Sur l'estimation non paramétrique des modèles conditionnels pour variables fonctionnelles spatialement dépendantes

Zoulikha Kaid

Université Charles De Gaulle Lille 3
En cotutelle avec

L'Université Djillali Liabes Sidi Bel Abbès, Algérie

Thèse

présentée en vue de l'obtention du grade de

Docteur

Discipline : Mathématiques Appliquées
Option : Statistique

Présentée par

Zoulikha Kaid

le 09 Décembre 2012

Sur l'estimation non paramétrique des modèles conditionnels pour variables fonctionnelles spatialement dépendantes

Composition du Jury

M. Attouch	Univ. Sidi Bel Abbes, Algérie	Rapporteur
S. Dabo-Niang	Univ. Lille 3, France	Directeur de thèse
R. Emilion	Univ. Orléans, France	Rapporteur
A. Gheriballah	Univ. Sidi Bel Abbes, Algérie	Examinateur
A. Laksaci	Univ. Sidi Bel Abbes, Algérie	Directeur de thèse
M. Rachdi	Univ. Grenoble 2, France	Examinateur

Dédicace
À tous ceux qui me sont proches.

Remerciements

En premier lieu, je tiens à exprimer mon profond sentiment de respect et de reconnaissance à mes directeurs de thèse les Professeurs **Sophie Dabo-Niang** et **Ali Laksaci** pour leur encadrement et leurs encouragements durant toute la période de la réalisation de ce travail. Je voudrais leur exprimer ma gratitude pour leur disponibilité, pour le temps qu'ils ont passé à m'aider, à améliorer et compléter mes travaux, pour leurs conseils avisés qui m'ont permis de mener à bien ce travail de thèse.

Je tiens à exprimer aussi ma reconnaissance à Monsieur le Professeur **Richard Emilion** pour avoir accepté d'être rapporteur de cette thèse. Je lui suis gré de l'intérêt qu'il a porté à mon travail.

Je souhaite également remercier Monsieur le Professeur **Mohamed Attouch** pour avoir accepté d'être rapporteur de cette thèse ainsi que pour l'attention qu'il a porté à mon travail

Pour la pertinence et la justesse de leur commentaires, remarques et suggestions, tant sur le fond que sur la forme, encore merci aux rapporteurs.

Pour l'intérêt qu'il a bien voulu accorder à mes recherches en acceptant de participer au jury, je tiens à exprimer mes vifs remerciements à Monsieur le Professeur **Mustapha Rachdi** .

Je voudrais aussi remercier vivement Monsieur le Professeur **Abdelkader Gheriballah** pour sa participation au jury et pour l'intérêt qu'il porte à mes travaux.

Un grand merci à tous les membres du laboratoire EQUIPPE de Lille 3 qui m'ont toujours accueilli avec chaleur et amitié pendant mes séjours à Lille.

Je tiens à remercier de manière plus générale tous les membres du Laboratoire de Mathématiques de l'université Djillali Liabès de Sidi Bel Abbès.

Enfin, j'adresse mes remerciements à tous ceux qui ont participé de près ou de loin à la bonne réalisation de la soutenance de cette thèse.

Table des matières

Chapitre 1

Introduction

1.1 Résumé

Dans cette thèse, nous nous intéressons au problème de la prévision spatiale en considérant des modèles non paramétriques conditionnels dont la variable explicative est fonctionnelle. Plus précisément, les points étudiés pour décrire la co-variation spatiale entre une variable réponse réelle et une variable fonctionnelle sont le mode conditionnel et les quantiles conditionnels.

En ce qui concerne le mode conditionnel, nous établissons la convergence presque complète, la convergence en norme L^p et la normalité asymptotique d'un estimateur à noyau. Ces propriétés asymptotiques sont obtenues sous des conditions assez générales telles, l'hypothèse de mélange forte et l'hypothèse de concentration de la mesure de probabilité de la variable explicative fonctionnelle. L'implémentation de l'estimateur construit en pratique est illustrée par une application sur des données météorologiques.

Le modèle des quantiles conditionnels est abordé dans la deuxième partie de la thèse. Il est traité comme fonction inverse de la fonction de répartition conditionnelle qui est estimée par un estimateur à double noyaux. Sous les mêmes conditions que celles du modèle précédent, nous donnons l'expression de la vitesse de convergence en norme L^p et nous démontrons la normalité asymptotique de l'estimateur construit.

Notre étude généralise au cas spatial de nombreux résultats déjà existant en série chronologique fonctionnelle. De plus, l'estimation de nos modèles repose sur une estimation préalable de la densité et de la fonction de répartition conditionnelles et permet de construire des régions prédictives, montrant ainsi l'apport de ce genre de modèles par rapport à la régression classique.

1.2 Summary

The main purpose of this thesis concerns the problem of spatial prediction using some nonparametric conditional models where the covariate variable is a functional one. More precisely, we treat the nonparametric estimation of the conditional mode and that of the conditional quantiles as spatial prediction tools alternative to the classical spatial regression of real response variable given a functional variable.

Concerning the first model, that is the conditional mode, it is estimated by maximizing the spatial version of the kernel estimate of the conditional density. Under a general mixing condition and the concentration properties of the probability measure of the functional variable, we establish the almost complete convergence (with rate), the L^p consistency (with rate) and the asymptotic normality of the considered estimator. The usefulness of this estimation is illustrated by an application on real meteorological data.

The model of the conditional quantiles is considered in the second part of this thesis and is treated as the inverse function of the conditional cumulative distribution function which is estimated by a double kernel estimator. Under the same general conditions as in the first model, we give the convergence rate in the L^p- norm and we show the asymptotic normality of the constructed estimator. These asymptotic results are closely related to the concentration properties on small balls of the probability measure of the underlying explanatory variable and the regularity of the conditional cumulative distribution function.

Our study generalizes to spatial case some existing results in functional times series case. Finally, we highlight what our models brings compared to classical regression, discussing the use of our results as preliminary works to construct predictive regions.

1.3 Statistique non paramétrique fonctionnelle : Motivation et contexte bibliographique

Ces deux dernières décennies, une immense innovation sur les appareils de mesures est apparue et réalisée permettant de surveiller plusieurs objets d'une manière continue, tels les indices boursiers, la pollution, la climatologie, les images satellitaires, ...
Ce développement technologique a exigé la modernisation des méthodes statistiques comme outils d'analyse et de contrôle. Ainsi, un nouvelle branche de la statistique, dénommée statistique fonctionnelle s'est développée pour traiter des observations comme éléments aléatoires fonctionnelles. Les premières contributions sur le sujet ont été consacrées à l'étude des modèles paramétriques (voir les monographes de Ramsay et Silverman (1997, 2002, 2005) pour le cas i.i.d ou Bosq (2000) pour le cas dépendant). Cependant, l'analyse statistique via les modèles linéaires est fondée sur une connaissance préliminaire de la nature de la co-variabilité entre les observations, ce qui est très difficile à vérifier en statistique fonctionnelle, contrairement à la statistique classique où on dispose d'outils graphiques comme le *scatterplot* qui

donne un aperçu sur la relation entre les observations. Ceci justifie l'importance de la modélisation des données fonctionnelles par des méthodes non paramétriques.
Le traitement non paramétrique des données fonctionnelles est beaucoup plus récent que l'analyse paramétrique. En effet, les premiers résultats ont été obtenus par Gasser et *al.* (1998). Ils se sont intéressés à l'estimation non paramétrique du mode de la distribution d'une variable fonctionnelle vérifiant un condition fractale. En considérant la même condition fractale Ferraty et Vieu (2000) ont étudié la convergence presque complète d'un estimateur à noyau de la fonction de régression, lorsque les observations sont indépendantes et identiquement distribuées. Dabo-Niang (2002) a obtenu, la convergence presque sûre et la normalité asymptotique d'un estimateur de type histogramme de la densité d'une variable aléatoire dans un espace de dimension infinie. En utilisant la propriété de concentration de la mesure de probabilité de la variable explicative fonctionnelle, Dabo-Niang et Rhomari (2004) ont étudié la convergence en norme L^p de l'estimateur à noyau de la régression non paramétrique. La convergence presque complète pour le cas fortement mélangeant a été étudié par Ferraty et *al.* (2004). Masry (2005) a montré la normalité asymptotique dans le cas d'observations fonctionnelles α-mélangeantes.
Les premiers résultats sur les modèles conditionnels ont été obtenus par Ferraty *et al.* (2006). Ils ont précisé la vitesse de convergence presque complète des estimateurs à noyau pour la fonction de répartition conditionnelle, la densité conditionnelle et ses dérivées, le mode conditionnel et les quantiles conditionnels. Nous renvoyons à Ferraty et Vieu (2006) pour un large éventail d'applications de ces modèles en statistique fonctionnelle. Dabo-Niang et Laksaci (2007) ont ajouté des résultats sur la convergence en norme L^p de l'estimateur à noyau du mode conditionnel dans le cas i.i.d. La détermination des termes dominants de l'erreur quadratique de l'estimateur à noyau de la densité conditionnelle a été obtenue par Laksaci (2007). Ferraty *et al* (2008) ont abordé l'estimation de la fonction du hasard conditionnelle et ont établi la convergence presque complète d'un estimateur à noyau de ce modèle non paramétrique. La normalité asymptotique des estimateurs à noyau du mode conditionnel et des quantiles conditionnels a été étudié par Ezzahrioui et Ould-said ont (2008a, 2008b, 2008c) en traitant deux cas (cas i.i.d et cas mixing). En considérant des observations α- mélangeantes, Quantela-del Rio (2008) a établi la convergence presque complète et la normalité asymptotique de l'estimateur proposé par Ferraty *et al* (2008) sur la fonction de hasard conditionnelle. L'auteur a illustré ces résultats asymptotiques par une application sur des données sismiques. Un estimateur alternatif des quantiles conditionnels a été proposé par Lemdani et al . (2009). Ils ont traité les quantiles conditionnels comme un modèle robuste appartenant à la classe des M-estimateurs. Les résultats asymptotiques de cet article sont la convergence presque complète et la normalité asymptotique dans le cas i.i.d.
La contribution de Ferraty et al.(2010) sur la convergence uniforme est très déterminante. Dabo-Niang en collaboration avec Laksaci (2010) ont généralisé leur résultats de la convergence en norme L^p de l'estimateur à noyau du mode conditionnel dans le cas i.i.d au cas fortement mélangeant. En considérant la même structure de dépendance, Lemdani et al . (2011) ont étudié la convergence presque complète et la normalité asymptotique de l'estimateur L^1 des quantiles conditionnels. Tandis que la convergence en norme L^p pour l'estimateur

à doubles noyaux des quantiles conditionnels a été récemment obtenue par Dabo-Niang et Laksaci (2012). La question du choix du paramètre de lissage dans l'estimation de la densité conditionnelle à variable explicative fonctionnelle a été considérée par Laksaci *et al.* (2012). D'autres auteurs se sont intéressés à l'estimation de la fonction de régression en utilisant d'autres approches, telles que la méthode des k plus proches voisins de Burba et al. (2008), les techniques robustes, voir Azzidine *et al.* (2008), Crambes et al. (2008), et Attouche et *al* . (2009, 2010). Pour l'estimation par la méthode des polynômes locaux, on peut voir Baìllo et Grané (2009) Barrientos-Marin et al. (2010), Berlinet et al. (2011) et Demongeot *et al.* (2012). La littérature sur le cas d'une variable réponse fonctionnelle est très restreinte en statistique fonctionnelle. On citera, dans ce contexte l'article de Dabo-Niang et Rhomari (2009) pour la convergence en norme L^p de l'estimateur à noyau de l'opérateur de régression comme élément banachique. La convergence presque complète de cet estimateur est obtenue par Ferraty et al. (2011). Van Keilegom en collaboration avec Ferraty et Vieu (2012) ont établi la normalité asymptotique de l'estimateur de la fonction de régression, lorsque les deux variables (réponse, explicative) sont de nature fonctionnelle. Tous ces résultats ont été obtenus dans le cas où les observations sont indépendantes identiquement distribuées. Le cas dépendant a été récemment considéré par Ferraty et al. (2012). Dans cette publication, les auteurs ont démontré la convergence presque complète de l'estimateur à noyau de l'opérateur de régression pour des observations β mélangeantes.

1.4 Les modèles conditionnels en statistique non paramétrique : Notes historiques et bibliographiques

L'étude de modèles non paramétriques liés à la distribution conditionnelle a été largement considérée en statistique non paramétrique. Historiquement, les premiers résultats sur ces modèles ont été obtenus par Roussas (1969). Il a traité l'estimation de la fonction de répartition conditionnelle par la méthode à noyau en utilisant des observations markoviennes. Il a établi la convergence en probabilité de l'estimateur construit. Un estimateur alternatif pour le même modèle a été élaboré par Stone (1977). Ce dernier a étudié l'estimateur empirique de la fonction de répartition conditionnelle et il a appliqué les résultats obtenus à l'estimation des quantiles conditionnelles comme inverse généralisé de la fonction de répartition conditionnelle. Stute (1986) a ajouté des résultats sur la convergence presque complète de l'estimateur à noyau de la fonction de répartition d'une variable aléatoire vectorielle conditionnellement à une variable explicative vectorielle. L'estimation du mode conditionnel fut traitée pour la première fois par Collomb et *al.* (1987). Ces auteurs ont montré la convergence uniforme de l'estimateur à noyau de ce modèle conditionnel lorsque les observations sont ϕ-mélangeantes. En 1989, Samanta a étudié la normalité asymptotique de l'estimateur à noyau de quantiles conditionnels lorsque les observations sont indépendantes et identiquement distribuées. Ce dernier en collaboration avec Thavaneswaran en 1990, ont obtenu la même propriété asymptotique pour un estimateur à noyau du mode conditionnel en considérant le cas i.i.d. Roussas (1991) a établi la convergence presque sûre d'un estimateur à noyau des quantiles conditionnels lorsque les observations sont issues d'un processus de Markov. La contribution de Youndjé (1993) sur l'estimation de la densité conditionnelle est déterminante. Il a abordé la question du choix du paramètre de lissage en considérant le deux cas indépendant et dépendant. Quintela et Vieu (1997) ont traité le mode conditionnel comme étant le point annulant la dérivée d'ordre 1 de la densité conditionnelle et ont construit un estimateur pour ce modèle en utilisant l'estimateur à noyau de la dérivée de la densité conditionnelle. Ould-said (1997) a étudié l'estimateur à noyau du mode conditionnel à partir des observations ergodiques. Nous renvoyons à Berlinet et *al.* (1998a), Louani et Ould-Said (1999) pour la convergence en loi de l'estimateur à noyau du mode conditionnel dans le cas $\alpha-$ mélangeant. L'article de Berlinet *et al.* (1998b) donne un théorème général de la normalité asymptotique des estimateurs des quantiles conditionnels, indépendamment de la corrélation des observations. Zhou et Liang (2000) ont utilisé l'approche L^1 pour construire un estimateur de la médiane conditionnelle en utilisant des observations α-mélangeantes. Ils ont montré la normalité asymptotique de cet estimateur. La convergence en norme L^p de l'estimateur à noyau de la densité conditionnelle d'un processus markovien stationnaire a été obtenu par Laksaci et Yousfate (2002). Ioannides et Matzner (2002) ont construit un estimateur pour le mode conditionnel, lorsque, les observations sont entachées d'erreurs. Dans cet article les auteurs se focalisent sur la convergence presque sûr de l'estimateur proposé. Tandis que sa normalité asymptotique a été démontrée par les mêmes auteurs dans Ioannides et Matzner (2004). Gannoun et *al.* (2003) ont abordé l'estimation des quantiles conditionnels par la méthode L^1, ils ont établi la convergence presque complète et la normalité asymptotique. En considérant le même modèle et la même méthode d'estimation, Lin et Li (2007) ont étudié la normalité asymptotique à partir des variables associées. D'autres auteurs se sont intéressés à l'estimation des modèles conditionnels à partir des observations censurées ou tronquées (voir, par exemple, Lemdani et al (2009) Liang et Uña-Álvarez,(2010, 2011) , Khardani et al. (2010, 2011 et 2012), Ould Saïd et Tatachak (2011) ou Ould Saïd et Djabrane (2011)).

1.5 Statistique non paramétrique spatiale : Etat de l'art

L'analyse statistique des données spatiales connaît un intérêt grandissant en statistique mathématique. L'importance de ce sujet est motivée par le nombre important de problèmes concrets, pour lesquels les données sont de structure de dépendance spatiale. On trouvera des exemples d'applications dans différents domaines tels, l'économie (les courbes de consommation d'un produit quelconque dans différents payes), les sciences de l'environnement (les courbes de concentration d'un gaz polluant dans différentes régions) ou en agronomie (les courbes de pluviométrie dans différentes localités). On trouve aussi des exemples en foresterie, en prospection minière, en analyse radiographique... Des travaux précurseurs en statistique spatiale sont ceux de Cressie (1991), Ripley (1981), Guyon (1995), Anselin et Florax (1995). En ce qui concerne la modélisation non-paramétrique des données spatiales, nous renvoyons à Tran (1990) qui traite l'estimation de la densité par la méthode du noyau. Ce dernier a établi la convergence en probabilité et la normalité asymptotique de l'estimateur à noyau de la densité d'une variable aléatoire. L'estimation de ce modèle non paramétrique spatial par la méthode des k- plus proches voisins a été traité par Tran et Yakowitz (1993). Ces auteurs ont démontré la normalité asymptotique de l'estimateur construit en considérant un champ aléatoire fortement mélangeant. Carbon et al.(1996) ont élaboré des résultats sur la convergence en norme L^1 de l'estimateur à noyau de la densité d'un vecteur aléatoire. La convergence en norme L^∞ de cet estimateur est obtenue par Carbon et al.(1997). La fonction de régression a été considérée par Lu et Chen (2004). Ils ont étudié la convergence en probabilité de l'estimateur à noyau de la fonction de régression. La convergence presque complète et la normalité asymptotique de cet estimateur ont été obtenues par Biau et Cadre (2004). Hallin, Lu et Tran (2004) ont utilisé la méthode des polynômes locaux pour construire un estimateur de la régression spatiale ainsi que sa dérivée d'ordre 1. Comme résultat asymptotique, ces auteurs ont démontré la normalité asymptotique des estimateurs construits. Carbon et al. (2007) ont ajouté des résultats sur la régression spatiale via le modèle autorégressif. Certain auteurs ont aussi publié des travaux (voir Li et Tran (2007)) traitant la normalité asymptotique de l'estimateur à noyau de la fonction de hasard. Xu et Wang (2008) ont étudié l'estimation locale linéaire de la régression spatiale en minimisant le critère de l'erreur absolue. La version spatiale de l'estimateur de la fonction de régression par la méthode des k-plus proches voisins a été proposée par Li et Tran (2009). Dans cet article, les auteurs ont montré la normalité asymptotique de l'estimateur considéré. L'estimation de la régression robuste en statistique non paramétrique spatiale a été abordée par Gheriballah et al. (2010). La plupart des articles sur-cités ont établi des propriétés asymptotiques de modèles spatiaux étudiés sous des conditions de mélange forte. Récemment, Robinson (2011) a établi la convergence en probabilité et la normalité asymptotique de l'estimateur à noyau de la régression sous des conditions moins restrictives. Notons aussi El-Mechkouri (2011) a affaibli les conditions de la normalité asymptotique de l'estimateur de Tran (1990). Ce dernier a élaboré une nouvelle méthode de démonstration ne nécessitant pas l'utilisation de la décomposition spatiale de Tran (1990). Les modèles non paramétriques conditionnels ont aussi fait l'objet de plusieurs publications en statistique spatiale. A ce sujet, on citera la contribution de Hallin et al. (2009) pour l'estimation locale linéaire de quantiles conditionnels. L'estimation par la méthode du noyau pour le mode et les quantiles conditionnels a été abordé par Ould Abdi et al. (2010a, 2010b). Ces auteurs se sont intéressés à la convergence en norme Lp et la normalité asymptotique de ces estimateurs. Récemment, Dabo-Niang et Thiam (2010) ont proposé un estimateur alternatif pour les quantiles conditionnels sont vus comme un modèle robuste. La convergence presque complète ainsi que la normalité asymptotique de cet estimateur ont été établis.

En ce qui concerne la modélisation fonctionnelle spatiale, qui est le sujet de ce présent travail, les premiers résultats disponibles furent établis par Basse et al. (2008). Ils ont étudié la version spatiale de l'estimateur de la densité d'une variable aléatoire spatiale prenant ses valeurs dans un espace fonctionnel. L'estimation du mode (non conditionnel) a été traitée et implémentée sur des données réelles par Dabo-Niang et al. (2009). Dans la même année, Laksaci et Maref (2009) ont donné la vitesse de convergence presque complète de l'estimateur à double noyaux de la fonction quantile. Laksaci et Mechab (2010) ont généralisé le travail de Ferraty et al. (2008) sur la fonction de hasard conditionnelle dans le cas i.i.d. au cas de données spatialement dépendantes. L'estimation de la fonction de régression dans ce contexte fonctionnel spatial a été considérée par Dabo-Niang et al. (2011). Dans cet article, les auteurs donnent l'expression de la vitesse de convergence presque sûre de la version spatiale de l'estimateur à noyau de la régression non paramétrique fonctionnelle. La normalité asymptotique de la régression robuste en statistique non paramétrique fonctionnelle spatiale a été obtenue par Attouch et al. (2012a). Tandis que sa convergence presque complète a été établie par Attouch et al. (2012b). L'estimation non paramétrique de la régression spatiale par la méthode locale linéaire a été récemment abordée par Chouaf et Laksaci (2012). Ils ont démontré la convergence presque complète de la version spatiale de l'estimateur de Barrientos-Marin et al. (2010). Nous renvoyons aux contributions de Biau (2003) et Dabo-Niang et Yao (2007) pour l'estimation de la densité et la fonction de régression à partir d'un champ aléatoire dont les indices sont continus. Nous renvoyons également à Giraldo (2009) et Nerini al. (2010) pour l'étude de modèles linéaires et ces applications en statistique fonctionnelle spatiale.

1.6 Contribution de la thèse

Dans cette contribution on se propose d'étudier l'estimation de certaines modèles conditionnels dans le cadre de la statistique non paramétrique fonctionnelle spatiale. Plus précisément nous considérons le mode conditionnel et les quantiles conditionnels. Le mode est estimé pour le point qui maximise l'estimateur à noyau de la densité conditionnelle, tandis que l'estimation du quantiles est obtenue par inversion de l'estimateur à noyau de la fonction de répartition conditionnelle.
Nous démontrerons la convergence presque complète, la convergence en norme L_p et la normalité asymptotique de nos estimateurs. Les résultats obtenus généralisent au cas spatial les travaux de Ferraty et al. (2005), Dabo-Niang et Laksaci (2010, 2012) sur l'estimation du mode et des quantiles conditionnels en série chronologique fonctionnelle.
En termes d'application, cette contribution a une grande importance en pratique. En effet, il s'agit d'adopter dans le cas spatial deux outils de prévision alternatives à la régression classique. De plus, ces deux méthodes de prévision peuvent être beaucoup plus robustes que celle de la régression classique qui est très sensible aux erreurs d'observations telle que la présence des observations aberrantes, et l'hétéroscedasticticité des données, l'asymétrie, la bi-modalité de la distribution ....
Notons qu'en pratique les variables fonctionnelles ne sont pas totalement observées, on observe seulement un nombre fini de points et on reconstitue l'observation fonctionnelle à l'aide d'une base considérée dans l'espace fonctionnel telle que les fonctions splines, les ondelettes, la base de Fourier, ... Ainsi, ce pré-traitement des observations fonctionnelles est étroitement lié à la base utilisée et à la grille de discrétisation de la trajectoire de la variable fonctionnelle, ce qui rend ces variables exposées aux problèmes mentionnés ci-dessus. En effet, ce pré-traitement est souvent confronté à des erreurs de mesure, d'enregistrement, de

problèmes de compatibilité avec la base,.... Il est évident que cette difficulté est multipliée en statistique fonctionnelle spatiale du moment que les sites d'observation (trajectoire en série temporelle) sont multidimensionnelles. D'où l'importance d'utiliser dans le contexte de la statistique fonctionnelle spatiale des outils d'analyse robustes, tel le mode conditionnel, la médiane conditionnelle ou les quantiles conditionnels.

Rappelons que ces modèles conditionnels ont aussi d'autres avantages par rapport à la régression classique. En effet, ils offrent la possibilité de déterminer des régions prédictives (voir par exemple De Gooijer et Gannoun (2000)) qui sont beaucoup plus intéressantes que la prévision ponctuelle en pratique. Ces modèles sont aussi très utilisés en finance et/ou en assurance pour modéliser les risques de valeurs extrêmes.

1.7 Le plan de la thèse

Cette thèse est présentée en cinq chapitres. Dans le reste de ce premier chapitre, on liste en bref les notations utilisées ainsi que les résultats asymptotiques obtenus

Dans le deuxième chapitre, on considère le problème de l'estimation du mode conditionnel, dont on traite sa convergence presque complète. Sous des conditions assez générales, on construit un estimateur à noyau pour le mode conditionnel et on explicite sa vitesse de convergence presque complète. On discutera aussi dans ce chapitre de l'applicabilité de ce modèle aux problèmes de prévision à partir d'un champ aléatoire à temps continue.

Le troisième chapitre est consacré à l'étude de la convergence en norme L^p et à la normalité asymptotique de l'estimateur à noyau du mode conditionnel construit dans le deuxième chapitre. On précise sa vitesse de convergence en norme L^p et on démontre sa convergence en loi vers la loi normale. On mettra l'accent, aussi sur l'utilisation de ces propriétés asymptotiques pour résoudre quelques questions importantes pour l'implémentation pratique de notre estimateur. Plus précisément, on discutera de la détermination asymptotique exacte de l'erreur quadratique, la question du choix du paramètre de lissage et la détermination des intervalles de confiance seront aussi abordées. Un exemple d'application sur des données réelles est traité dans la fin de ce chapitre.

Le quatrième chapitre concerne l'estimation des quantiles conditionnels. On suppose qu'on dispose d'un champ aléatoire fonctionnel fortement mélangeant comme dans les chapitres précédents et on établit la convergence en norme L^p et la normalité asymptotique de l'estimateur à noyau de la fonction de quantile définie comme inverse de la fonction de répartition conditionnelle.

Pour donner une idée de quelques types d'application du cadre de nos travaux, des exemples d'applications ainsi qu'une discussion sont donnés dans le chapitre 5. Le dernier chapitre de cette contribution est consacré à quelques commentaires et discussions sur les nombreuses questions ouvertes qui en découlent.

1.8 Brève présentation des estimateurs étudiés dans la thèse

Soit N un entier naturel dans $\mathbb{N}^*$. On considère le champ aléatoire $Z_{\mathbf{i}} = (X_{\mathbf{i}}, Y_{\mathbf{i}})$, $\mathbf{i} \in \mathbb{N}^N$ à valeurs dans $\mathcal{F} \times \mathbb{R}$, où $(\mathcal{F}, d)$ est un espace semi-métrique de dimension éventuellement infinie. Pour tout $x \in \mathcal{F}$ on définit la fonction de répartition conditionnelle de Y sachant $X = x$, notée F^x par

$$F^x(y) = P(Y \leq y | X = x), \quad \forall y \in \mathbb{R}.$$

On suppose que cette distribution conditionnelle de Y sachant X est absolument continue par rapport à la measure de Lebesgue dont, on désigne par f^x la densité conditionnelle et par $f^{x(j)}$ la dérivée d'ordre j de cette densité conditionnelle.

Par ailleurs, nous supposons que le champ aléatoire fonctionnel est observé sur l'ensemble $I_{\mathbf{n}} = \{\mathbf{i} = (i_1, ..., i_N) \in \mathbb{N}^N\ ,\ 1 \leq i_k \leq n_k,\ k = 1, ..., N\}$, $\mathbf{n} = (n_1, ..., n_N) \in \mathbb{N}^N$ et on estime la fonction de répartition conditionnelle par

$$\widehat{F}^x_{\mathbf{n}}(y) = \begin{cases} \dfrac{\sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K_1\left(\frac{d(x, X_{\mathbf{i}})}{a_{\mathbf{n}}}\right) K_2\left(\frac{y - Y_{\mathbf{i}}}{b_{\mathbf{n}}}\right)}{\sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K_1\left(\frac{d(x, X_{\mathbf{i}})}{a_{\mathbf{n}}}\right)} & \text{si} \quad \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K_1\left(\frac{d(x, X_{\mathbf{i}})}{a_{\mathbf{n}}}\right) \neq 0 \\ 0 & \text{sinon} \end{cases}.$$

où K_1 est un noyau, K_2 une fonction de répartition et $a_{\mathbf{n}}$ et $b_{\mathbf{n}}$ sont des suites de nombres réels positifs.
De $\widehat{F}^x_{\mathbf{n}}$, on déduit un estimateur de la densité conditionnelle, noté $\widehat{f}^x$, défini par

$$\widehat{f}^x(y) = \frac{h_H^{-1} \sum_{\mathbf{i} \in I_{\mathbf{n}}} K(h_K^{-1} d(x, X_{\mathbf{i}})) H(h_H^{-1}(y - Y_{\mathbf{i}}))}{\sum_{\mathbf{i} \in I_{\mathbf{n}}} K(h_K^{-1} d(x, X_{\mathbf{i}}))}, \qquad \forall y \in \mathbb{R}$$

où $H = K_2'$ est la dérivée de K_2 et $h_K = h_{K,n} = a_{\mathbf{n}}$ (respectivement, $h_H = h_{H,n} = b_{\mathbf{n}}$).
Pour l'estimation du mode conditionnel, on suppose qu'il existe un compact S où le mode est unique. On note par $\theta(x)$ ce mode. L'estimateur de ce dernier est obtenu par la maximisation de l'estimateur de la densité conditionnelle. Plus précisément, on estime $\theta(x)$ par $\hat{\theta}(x)$ vérifiant

$$\widehat{f}^x(\hat{\theta}(x)) = \sup_{y \in S} \widehat{f}^x(y).$$

En ce qui concerne l'estimation du quantile conditionnel d'ordre $\alpha \in]0, 1[$, on note $q_\alpha(x)$ ce quantile, solution de l'équation

$$F^x(q_\alpha(x)) = \alpha.$$

Pour assurer l'existence et l'unicité de $q_\alpha(x)$ on suppose que la fonction de répartition conditionnelle F^x est strictement croissante. On estime le quantile conditionnel $q_\alpha(x)$ par $\hat{q}_\alpha(x)$ tel que

$$\widehat{F}^x_{\mathbf{n}}(\hat{q}_\alpha(x)) = \alpha.$$

1.9 Brève présentation des résultats de convergence obtenus dans la thèse

Dans cette these, on suppose que le champ aléatoire fonctionnel $(Z_{\mathbf{i}})_{\mathbf{i}\in\mathbb{N}^N}$ vérifie la condition de mélange suivante

$$\left\{\begin{array}{c} \text{Il existe une fonction } \varphi\left(t\right)\downarrow 0 \text{ quand } t\rightarrow\infty, \text{telle que} \\ \forall\, E,\ E^{'} \text{sous ensembles de } \mathbb{N}^N \text{de cardinal fini} \\ \alpha\left(\mathcal{B}(E),\,\mathcal{B}(E^{'})\right)=\sup_{B\in\mathcal{B}(E),\,C\in\mathcal{B}(E^{'})}\left|\mathbb{P}\left(B\cap C\right)-\mathbb{P}\left(B\right)\mathbb{P}\left(C\right)\right| \\ \leq\psi\left(\text{Card}\left(E\right),\text{Card}(E^{'})\right)\varphi\left(\text{dist}\left(E,E^{'}\right)\right), \end{array}\right.$$

où $\mathcal{B}(E)$ (*resp.* $\mathcal{B}(E^{'})$) la tribu borélienne engendrée par $(Z_{\mathbf{i}},\ \mathbf{i}\in E)$ (*resp.* $\left(Z_{\mathbf{i}},\ \mathbf{i}\in E^{'}\right)$), $\text{Card}(E)$ (*resp.* $\text{Card}(E^{'})$) est le cardinal de E (*resp.* $E^{'}$), $\text{dist}\big(E,E^{'}\big)$ désigne la distance euclidienne entre E et $E^{'}$ et ψ une fonction symétrique : $\mathbb{N}^2\rightarrow\mathbb{R}^+$, décroissante par rapport aux deux variables et satisfait la condition suivante

$$\psi\left(n,m\right)\leq C\min\left(n,m\right),\qquad\forall n,m\in\mathbb{N}. \tag{1}$$

Par ailleurs, on suppose que nos modèles vérifient quelques conditions de régularité et que les variables fonctionnelles $X_{\mathbf{i}}$ vérifient l'hypothèse de concentration. On obtient ainsi les résultats suivants.

Sur le mode conditionnel

Théorème 1

$$\widehat{\theta}(x)-\theta(x)=O\left(h_K^{\frac{b_1}{j}}+h_H^{\frac{b_2}{j}}\right)+O\left(\left(\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H\phi_x(h_K)}\right)^{\frac{1}{2j}}\right),\qquad\textit{presque complètement.}$$

où $\widehat{\mathbf{n}}=n_1...n_N$, *et* $\phi(h_K)$ *est la probabilité que la variable fonctionnelle* X *soit dans la boule de centre* x *et de rayon* h_K.

La preuve de ce résultat et le détail des conditions imposées serons donnés dans le Chapitre 2.

Théorème 2

$$\left\|\widehat{\theta}(x)-\theta(x)\right\|_p=O\left(h_K^{b_1}+h_H^{b_2}\right)+O\left(\left(\frac{1}{\widehat{\mathbf{n}}\,h_H^3\phi_x(h_K)}\right)^{\frac{1}{2}}\right)$$

où $\|.\|_p=(E^{1/p}|.|^p)$.

Théorème 3 *Pour tout $x \in \mathcal{A}$,*

$$\left(\frac{\widehat{\mathbf{n}}h_H^3\phi_x(h_K)}{\sigma^2(x)}\right)^{1/2}\left(\widehat{\theta}(x)-\theta(x)-B_{\mathbf{n}}(x)\right)\xrightarrow{\mathcal{D}}\mathcal{N}(0,1)\quad as\quad \mathbf{n}\to\infty.$$

où

$$B_{\mathbf{n}}(x)=B_H(x)h_H+B_K(x)h_K$$

avec

$$\begin{aligned} B_H(x) &= \frac{1}{2}\frac{\partial^2 f^x(\theta(x))}{\partial y^2}\int tH(t)dt \\ B_K(x) &= h_K\Phi_0'(0)\frac{\left(K(1)-\int_0^1(sK(s))'\beta_x(s)ds\right)}{\left(K(1)-\int_0^1 K'(s)\beta_x(s)ds\right)}. \end{aligned}$$

et

$$\sigma^2(x)=\frac{\beta_2 f^x(\theta(x))}{\beta_1^2}\int H'^2(t)dt\ \ (avec\ \beta_j=K^j(1)-\int_0^1(K^j)'(s)\beta_x(s)ds,\ for,\ j=1,\,2),$$

où

$$\mathcal{A}=\{x\in\mathcal{F},\ f^x(y)(1-F^x(y))\neq 0\}.$$

Les démonstrations de ces deux derniers résultats et le détail des conditions imposées serons donnés au Chapitre 3.

Sur le quantile conditionnel

Le résultats suivants concernent respectivement la convergence en norme Lp et la normalité asymptotique de l'estimateur par quantile conditionnel dans un carde spatial.

Théorème 4

$$\|\hat{q}_\alpha(x)-q_\alpha(x)\|_p=\left(E\left|\hat{q}_\alpha(x)-q_\alpha(x)\right|^p\right)^{1/p}=O\left((a_{\mathbf{n}})^{b_1}+(b_{\mathbf{n}})^{b_2}\right)+O\left(\left(\frac{1}{\widehat{\mathbf{n}}\,\phi_x(a_{\mathbf{n}})}\right)^{\frac{1}{2}}\right).$$

Théorème 5 *Pour tout $x \in \mathcal{A}$,*

$$\left(\frac{(f^x(q_\alpha(x)))^2\widehat{\mathbf{n}}(\psi_{K_1}(a_{\mathbf{n}}))^2}{\psi_{K_1^2}(a_{\mathbf{n}})(\alpha(1-\alpha))}\right)^{(1/2)}\left(\widehat{q_\alpha}(x)-q_\alpha(x)-C_{\mathbf{n}}(x)\right)\to_{n\to+\infty}N(0,1),$$

où

$$C_{\mathbf{n}}(x)=\frac{1}{f^x(q_\alpha(x))\psi_{K_1}(a_{\mathbf{n}})}\left(\alpha\psi_{K_1}(a_{\mathbf{n}})-E\left[K_1\left((a_{\mathbf{n}})^{-1}d(x,X)\right)F^X(q_\alpha(x))\right]\right)+O(b_{\mathbf{n}})$$

avec

$$\mathcal{A}=\{x\in\mathcal{F},\ \frac{\psi_{K_1^2}(a_{\mathbf{n}})}{(\psi_{K_1}(a_{\mathbf{n}}))^2}\neq 0\}\ avec\ \psi_g(h)=-\int_0^1 g'(t)\phi_x(ht)dt.$$

Les hypothèses imposées et les preuves des résultats ci-dessus seront données au chapitre 4.

Chapitre 2

La convergence presque complète de l'estimateur spatial du mode conditionnel

L'objectif de ce chapitre est de montrer la convergence presque complète de l'estimateur à noyau du mode conditionnel lorsque les observations sont fonctionnelles et spatialement dépendantes. Comme toute étude asymptotique en statistique non paramétrique fonctionnelle la difficulté technique se situe dans le fait de la non-existence de la mesure de Lebesgue dans les espaces de dimension infinie. Ainsi, l'apport principal de ce chapitre porte sur l'adaptation des outils classiques de la statistique non-paramétrique spatiale dans ce contexte fonctionnel où la notion d'intégration est plus difficile à utiliser. Cette étude met en évidence la propriété de concentration de la mesure de probabilité de la variable fonctionnelle considérée sur les petites boules. *Les résultats de ce chapitre font l'objet d'une publication dans la revue "Statistics and Probability Letters".*

On spatial conditional mode estimation for a functional regressor

Sophie Dabo-Niang ♯*, Zoulikha Kaid ♯♮
Labo. EQUIPPE, Maison de la recherche
♯Université Lille 3,
BP 60149, 59653,Villeneuve d'Ascq cedex, France.
e-mail: sophie.dabo@univ-lille3.fr, zoulikha.kaid@etu.univ-lille3.fr

Ali Laksaci ♮
♮ Agence Nationale de Dveloppement
et de Recherche Universitaire (A.N.D.R.U.)
Laboratoire de Mathématiques
Univ. Djillali Liabès
BP 89, 22000 Sidi Bel Abbès, Algeria
e-mail: alilak@yahoo.fr

August 23, 2012

Abstract
Let $(Z_{\mathbf{i}} = (X_{\mathbf{i}}, Y_{\mathbf{i}}),\ \mathbf{i} \in \mathbb{Z}^N)$ be a $\mathcal{F} \times \mathbb{R}$-valued measurable strictly stationary spatial process, where $\mathcal{F}$ is a semi-metric space. We study a kernel estimator of the conditional mode of the univariate response variable $Y_{\mathbf{i}}$ given the functional variable $X_{\mathbf{i}}$. The main aim of this paper is to prove the almost complete convergence (with rate) of this estimate.

Keywords: Spatial process, conditional mode estimate, non-parametric, Functional data.

*corresponding author

2.1 Introduction

This paper deals with nonparametric modeling of functional data presenting spatial dependence. More precisely, we are interested in the spatial prediction problem via the conditional mode estimation with functional regressor. Note that the modelization of this kind of data has been selected by Ramsay (2008) among the eight most interesting research subjects in functional data analysis. This is motivated by the increasing number of situations coming from different fields of applied sciences for which the data are of functional nature and showing a spatial interaction. In the finite dimensional framework, the modelization of spatial data has received a lot of attention (see Haining (2003) for a list of references). The nonparametric treatment of such data has also, been widely developed in the last two decades. Key references on this subject are Tran (1990), Biau and Cadre (2004), Dabo-Niang and Yao (2007), Li and Tran (2009), Dabo-Niang and Thiam (2010). However, there are few results in nonparametric modelization for functional spatial data. Specifically, Dabo-Niang *et al* . (2011a) stated the weak and the strong convergence of the functional spatial kernel estimate of the regression function. While, Laksaci and Maref (2009) obtained a rate of almost complete convergence of the kernel estimate of the spatial conditional quantiles where the regressor is of functional nature. The asymptotic normality of this estimate has been stated by Dabo-Niang *et al.* (2011b). Among the recent papers on the conditional mode estimation in functional statistics, we refer to Ferraty and Vieu (2006), Dabo-Niang and Laksaci (2010), Ezzahrioui and Ould-Said (2008) and Ferraty *et al.* (2010).
In this paper, we consider the problem of spatial prediction via a conditional mode estimation. Under some general mixing conditions, we establish almost complete convergence with rates of a kernel mode estimate. It should be noted that, the interest of our study of this context, comes mainly from the fact that the modal regression provides better prediction tool than the classical mean regression in several situations (see for instance Collomb *et al.* 1987, for some comparison examples between both models). The applicability of our study in a spatial prediction context has been discussed in this work.
The paper is organized as follows. We present our model in Section 2. In Section 3, we give some notations, hypotheses and the main results. Section 4 is devoted to some applications concerning continuously indexed spatial processes. The proofs of the results are given in the last Section.

2.2 The spatial conditional mode and its estimator

Let $(Z_{\mathbf{i}} = (X_{\mathbf{i}}, Y_{\mathbf{i}}),\ \mathbf{i} \in \mathbb{Z}^N)$ be a $\mathcal{F} \times \mathbb{R}$-valued measurable strictly stationary spatial process, defined on a probability space $(\Omega,\ \mathcal{A}, \mathbb{P})$, where $\mathcal{F}$ is a semi-metric space. Let d denote the semi-metric and $N \geq 1$. A point $\mathbf{i} = (i_1, ..., i_N) \in \mathbb{Z}^N$ will be referred to as a site. We assume that the process under study $(Z_{\mathbf{i}})_{\mathbf{i}}$ is observed over a rectangular domain $I_{\mathbf{n}} = \{\mathbf{i} = (i_1, ..., i_N) \in \mathbb{Z}^N\ ,\ 1 \leq i_k \leq n_k,\ k = 1, ..., N\}$, $\mathbf{n} = (n_1, ..., n_N) \in \mathbb{Z}^N$. A point $\mathbf{i}$ will be referred to as a *site*. We will write $\mathbf{n} \to \infty$ if $\min\{n_k\} \to \infty$ and $|\frac{n_j}{n_k}| < C$ for a constant C

such that $0 < C < \infty$ for all $j,\ k$ such that $1 \leq j, k \leq N$. For $\mathbf{n} = (n_1, ..., n_N) \in \mathbb{Z}^N$, we set $\widehat{\mathbf{n}} = n_1 \times \cdots \times n_N$.

In this work, we will assume that the functional random filed $(Z_{\mathbf{i}}, \mathbf{i} \in \mathbb{Z}^N)$ satisfies the following mixing condition :

$$(1) \qquad \begin{cases} \text{There exists a function } \varphi(t) \downarrow 0 \text{ as } t \to \infty, \text{ such that} \\ \forall\, E,\ E' \text{ subsets of } \mathbb{Z}^N \text{ with finite cardinals} \\ \alpha\left(\mathcal{B}(E), \mathcal{B}(E')\right) = \sup\limits_{B \in \mathcal{B}(E),\, C \in \mathcal{B}(E')} \left|\mathbb{P}(B \cap C) - \mathbb{P}(B)\mathbb{P}(C)\right| \\ \leq \psi\left(\operatorname{Card}(E), \operatorname{Card}(E')\right) \varphi\left(\operatorname{dist}(E, E')\right), \end{cases}$$

where $\mathcal{B}(E)$(*resp.* $\mathcal{B}(E')$) denotes the Borel σ-field generated by $(Z_{\mathbf{i}},\ \mathbf{i} \in E)$ (*resp.* $(Z_{\mathbf{i}},\ \mathbf{i} \in E')$), $\operatorname{Card}(E)$ (*resp.* $\operatorname{Card}(E')$) the cardinality of E (*resp.* E'), $\operatorname{dist}(E, E')$ the Euclidean distance between E and E' and $\psi : \mathbb{Z}^2 \to \mathbb{R}^+$ is a symmetric positive function nondecreasing in each variable such that

$$(2) \qquad \psi(n, m) \leq C \min(n, m), \qquad \forall n, m \in \mathbb{Z}$$

for some $C > 0$. We assume also that the process satisfies the following mixing condition :

$$(3) \qquad \sum_{i=1}^{\infty} i^{\delta} \varphi(i) < \infty, \ \ \delta > 0.$$

Note that conditions (4)-(6) are the same as the mixing conditions used by Tran (1990), and are satisfied by many spatial models (see Guyon (1987) for some examples). If $N = 1$, then $Z_{\mathbf{i}}$ is called strongly mixing time series.

Throughout the paper, we fix a point $x \in \mathcal{F}$ and we denote by $\mathcal{N}_x$ fixed neighborhood of x. Assume that the $Z_{\mathbf{i}}$'s have the same distribution as (X, Y) and there exists a regular version of the conditional probability of Y given X. Let F^x be the conditional distribution of the variable Y given $X = x$. Moreover, we suppose that F^x has a continuous density f^x with respect to (w.r.t.) Lebesgue's measure over $\mathbb{R}$ and we assume that there is some compact subset $S := [\alpha_x, \beta_x]$, such that the conditional density f^x has an unique mode $\theta(x)$ on S. A natural and usual estimator of $\theta(x)$ denoted $\widehat{\theta}(x)$, is given by :

$$(4) \qquad \widehat{\theta}(x) \ = \ \arg\sup_{y \in S} \widehat{f^x}(y),$$

where $\widehat{f^x}(\cdot)$ is the estimator of $f^x(\cdot)$ defined by

$$\widehat{f^x}(y) = \frac{h_H^{-1} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K(h_K^{-1} d(x, X_{\mathbf{i}})) H(h_H^{-1}(y - Y_{\mathbf{i}}))}{\sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K(h_K^{-1} d(x, X_{\mathbf{i}}))}, \qquad \forall y \in \mathbb{R}$$

with K and H are kernels functions and $h_K = h_{K,n}$ (resp. $h_H = h_{H,n}$) is a sequence of positive real numbers.

Note that this estimate $\widehat{\theta}(x)$ is not necessarily unique, so, the remainder of the paper concerns any value $\widehat{\theta}(x)$ satisfying (4.2).
The main goal of this paper is to study the nonparametric estimate $\widehat{\theta}(x)$ of $\theta(x)$ when the explanatory variable X is valued in the space $\mathcal{F}$ of eventually infinite dimension. Recall that, these questions in infinite dimension are particularly interesting, not only for the fundamental problems they formulate, but also for many applications, see Bosq (2000), Ramsay and Silverman (2005) and Ferraty and Vieu (2006).

2.3 Main results

All along the paper, when no confusion will be possible, we will denote by C or C' some strictly positive generic constants, and we will denote for all $\mathbf{i} \in \mathcal{I}_{\mathbf{n}}$, $K_{\mathbf{i}}(x) = K(h_K^{-1}d(x, X_{\mathbf{i}}))$, $H_{\mathbf{i}}(y) = H(h_H^{-1}(y - Y_{\mathbf{i}}))$ and $B(x,h) = \{x' \in \mathcal{F} \ : d(x',x) < h\}$. We shall list some required conditions that are necessary in deriving the almost complete convergence of the kernel estimate $\widehat{\theta}(x)$ of $\theta(x)$.

(H1) $P(X \in B(x,r)) = \phi_x(r) > 0$.

(H2) $\forall \mathbf{i} \neq \mathbf{j}$, $0 < \sup_{\mathbf{i}\neq\mathbf{j}} \mathbb{P}\left[(X_{\mathbf{i}}, X_{\mathbf{j}}) \in B(x,h_K) \times B(x,h_K)\right] \leq C(\phi_x(h_K))^{(a+1)/a}$, for $1 < a < \delta N^{-1}$.

(H3) $\forall (y_1,y_2) \in S \times S$, $\forall (x_1,x_2) \in N_x \times N_x$,

$$|f^{x_1}(y_1) - f^{x_2}(y_2)| \leq C\left(d(x_1,x_2)^{b_1} + |y_1 - y_2|^{b_2}\right), \quad b_1 > 0,\, b_2 > 0.$$

(H4) f^x is j-times continuously differentiable with respect to y on S such that, $f^{x(l)}(\theta(x)) = 0$, for $1 \leq l < j$, and $0 < \left|f^{x(j)}(y)\right| < \infty$, for all $y \in S$.

(H5) K is a function with support $(0,1)$ such that $0 < C' < K(t) < C < \infty$.

(H6) The kernel H such that

$$\forall (y_1,y_2) \in \mathbb{R}^2 \ |H(y_1) - H(y_2)| \leq C|y_1 - y_2|, \quad \int H(t)dt = 1 \quad \text{and} \quad \int |t|^{b_2} H(t)dt < \infty.$$

(H7) There exists $0 < \alpha < (\delta - 5N)/3N$ and $\eta_0 > 0$, such that $\lim_{\mathbf{n}\to\infty} \widehat{\mathbf{n}}^{\alpha}\, h_H = \infty$ and $C\widehat{\mathbf{n}}^{\frac{(5+3\alpha)N-\delta}{\delta}+\eta_0} \leq h_H\phi_x(h_K)$.

Notice that these conditions are very standard in this context. Indeed, condition (H1) is the same as that used by Ferraty and Vieu (2006). It is a condition closely linked with the semi-metric d. So, a right choice of d can provide a solution to the curse of dimensionality. Moreover, the function $\phi_x(.)$ defined in (H1) can be explicitly given for several continuous processes. The local dependence (H2) allows to get the same convergence rate as in the i.i.d. case (see Ferraty and Vieu (2006)). Assumptions (H3) and (H4) are regularity conditions which characterize the functional space of our model and are needed to evaluate the bias term in the asymptotic results of this paper. The hypotheses (H5)-(H7) are technical conditions and are also similar to those considered in Ferraty *et al.* (2005).

The following Theorem gives the almost-complete convergence[1] (a.co.) of $\widehat{\theta}(x)$.

Theorem 2.3.1 *Under assumptions (H1)-(H7), we have :*

$$\widehat{\theta}(x)-\theta(x)=O\left(h_K^{\frac{b_1}{j}}+h_H^{\frac{b_2}{j}}\right)+O\left(\left(\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H\phi_x(h_K)}\right)^{\frac{1}{2j}}\right),\qquad a.co.$$

where $\widehat{\mathbf{n}}=n_1...n_N$, *$j$ is given in (H4).*

Remark 1
Here, we point out that, if $N=1$, we obtain the same rate as in i.i.d. case given in Ferraty and Vieu (2006).

The proof of Theorem 2.3.1 is based on the following results.

Lemma 2.3.2 *Under the hypotheses (H1)-(H2), (H5) and (H7), we have*

$$\widehat{f}_D^x-\mathrm{E}\widehat{f}_D^x=O_{a.co.}\left(\left(\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,\phi_x(h_K)}\right)^{\frac{1}{2}}\right)\qquad and\qquad \sum_{\mathbf{n}}\mathbb{P}\left(\widehat{f}_D^x<1/2\right)<\infty,$$

where

$$\widehat{F}_D^x \;=\; \frac{1}{\widehat{\mathbf{n}}\mathrm{E}K_{\mathbf{1}}(x)}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}K\left(h_K^{-1}d(x,X_i)\right).$$

Lemma 2.3.3 *Under the hypotheses (H1) and (H4)-(H7), we have,*

$$\sup_{y\in S}\left|f^x(y)-\mathrm{E}\widehat{f}_N^x(y)\right|=O\left(h_K^{b_1}+h_H^{b_2}\right)$$

where $\widehat{f}_N^x(y)=\dfrac{1}{\widehat{\mathbf{n}}\,h_H\mathrm{E}K_{\mathbf{1}}(x)}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}K\left(h_K^{-1}d(x,X_{\mathbf{i}})\right)H\left(h_H^{-1}(y-Y_{\mathbf{i}})\right).$

Lemma 2.3.4 *Under the hypotheses of Theorem 2.3.1, we have,*

$$\sup_{y\in S}\left|\widehat{f}_N^x(y)-\mathrm{E}\widehat{f}_N^x(y)\right|=O_{a.co.}\left(\left(\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H\phi_x(h_K)}\right)^{\frac{1}{2}}\right).$$

1. Let $(z_n)_{n\in\mathbb{N}}$ be a sequence of real r.v.'s; we say that z_n converges almost completely (a.co.) to zero if, and only if, $\forall\epsilon>0$, $\sum_{n=1}^{\infty}P(|z_n|>\epsilon)<\infty$. Moreover, we say that the rate of almost complete convergence of z_n to zero is of order u_n (with $u_n\to0$) and we write $z_n=O_{a.co.}(u_n)$ if, and only if, $\exists\epsilon>0$, $\sum_{n=1}^{\infty}P(|z_n|>\epsilon u_n)<\infty$. This kind of convergence implies both almost sure convergence and convergence in probability.

2.4 Application to continuously indexed random fields

Clearly, a continuously indexed random filed $(Z_t)_{t\in\mathbb{R}^N}$ is one of important examples of functional spatial data. Indeed, let $(Z_t)_{t\in\mathbb{R}^N}$ be a $\mathbb{R}$-valued strictly stationary random spatial processus assumed to be bounded and observed over some subset $I \subset \mathbb{R}^N$. Then, our approach can be used to predict the value Z_{s_0} at an unobserved location $s_0 \notin I$ by taking into account, the observed part of $(Z_{t\in I})$ in its continuous form. For this, we suppose that, the value of Z_{s_0} depends only on the values of the process (Z_t) in a bounded neighborhood $\mathcal{V}_{s_0} \subset I$ of s_0. From Z_t we may construct m functional spatial random variables as follows : Consider some grid $\mathcal{G}_{\mathbf{n}} = \{t_i = (t_{i,1}, \dots t_{i,N}) \in I\ ,\ 1 \le t_{i,j} \le n_j,\ j = 1, ..., N,\ i = 1, ..., m\}$ such that

$$\forall i = 1, \dots m \qquad \min_{1\le j\le N-1}(t_{i,j+1} - t_{i,j}) \ge C > 0 \quad \text{for some constant } C$$

and we define

$$\forall i = 1, ..., m, \qquad X_{t_i} = (Z_t,\ t \in \mathcal{V}_{t_i})$$

where $\mathcal{V}_{t_i} = \mathcal{V} + t_{t_i}$, $\mathcal{V} = \mathcal{V}_{s_0} - s_0$, which does not contain 0. So, the predictor that we proposed (see Biau and Cadre (2004), Dabo-Niang and Yao (2007) for the finite dimension mean regression case), aims to evaluate a real characteristic denoted $Y_{s_0} = Z_{s_0}$, at a site s_0, given $X_{s_0} = (Z_t,\ t \in \mathcal{V}_{s_0})$. The random variable $\widehat{\theta}(X_{s_0})$, is in some situations, a good approximation of the quantity Y_{s_0}. The latter is given by using the m pairs of r.v $(X_{t_i}, Y_{t_i})_i$ with $Y_{t_i} = Z_{t_i}$.

2.5 Appendix

We state the following lemmas.

Lemma 2.5.1 *(See Carbon et al (1997)) Suppose $E_1, ..., E_r$ are sets containing m sites each with $dist(E_i,\ E_j) \ge \gamma$ for all $i \ne j$ where $1 \le i \le r$ and $1 \le j \le r$. Suppose $Z_1, ..., Z_r$ is a sequence of real-valued r.v.'s measurable with respect to $\mathcal{B}(E_1), ..., \mathcal{B}(E_r)$ respectively, and Z_i takes values in $[a, b]$. Then, there exists a sequence of independent r.v.'s $Z_1^*, ..., Z_r^*$ independent of $Z_1, ..., Z_r$ such that Z_i^* has the same distribution as Z_i and satisfies*

$$\sum_{i=1}^{r} \mathrm{E}|Z_i - Z_i^*| \le 2r(b-a)\psi((r-1)m, m)\varphi(\gamma).$$

Lemma 2.5.2 *(See Carbon et al (1997))*
(i) Suppose that (3) holds. Denote by $\mathcal{L}_r(\mathcal{F})$ the class of $\mathcal{F}-$measurable r.v.'s X satisfying $\|X\|_r = (\mathrm{E}|X|^r)^{1/r} < \infty$. Suppose $X \in \mathcal{L}_r(\mathcal{B}(E))$ and $Y \in \mathcal{L}_r(\mathcal{B}(E'))$. Assume also that $1 \le r,\ s,\ t < \infty$ and $r^{-1} + s^{-1} + t^{-1} = 1$. Then

$$(5) \qquad |\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]| \le C\|X\|_r\|Y\|_s\{\psi(Card(E), Card(E'))\varphi(dist(E, E'))\}^{1/t}.$$

(ii) For r.v.'s bounded with probability 1, the right-hand side of (9) can be replaced by $C\psi(Card(E), Card(E'))\varphi(dist(E, E'))$.

Proof of Lemma 4.4.6.
Let $\Delta_{\mathbf{i}}(x) = K_{\mathbf{i}}(x) - \mathrm{E}\left[K_{\mathbf{i}}(x)\right]$, then

$$\widehat{f}_D^x - \mathrm{E}[\widehat{f}_D^x] = \frac{1}{\widehat{\mathbf{n}}\mathrm{E}\left[K_{\mathbf{1}}(x)\right]} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \Delta_{\mathbf{i}}(x).$$

Consider the spatial block decomposition on the random variables $\Delta_{\mathbf{i}}(x)$ for a fixed integer $p_{\mathbf{n}}$, (depending on $\mathbf{n}$) as follows

$$U(1,\mathbf{n},x,\mathbf{j}) = \sum_{\substack{i_k=2j_kp_{\mathbf{n}}+1\\ k=1,\ldots,N}}^{2j_kp_{\mathbf{n}}+p_{\mathbf{n}}} \Delta_{\mathbf{i}}(x),$$

$$U(2,\mathbf{n},x,\mathbf{j}) = \sum_{\substack{i_k=2j_kp_{\mathbf{n}}+1\\ k=1,\ldots,N-1}}^{2j_kp_{\mathbf{n}}+p_{\mathbf{n}}} \quad \sum_{i_N=2j_Np_{\mathbf{n}}+p_{\mathbf{n}}+1}^{(j_N+1)p_{\mathbf{n}}} \Delta_{\mathbf{i}}(x),$$

$$U(3,\mathbf{n},x,\mathbf{j}) = \sum_{\substack{i_k=2j_kp_{\mathbf{n}}+1\\ k=1,\ldots,N-2}}^{2j_kp_{\mathbf{n}}+p_{\mathbf{n}}} \quad \sum_{i_{N-1}=2j_{N-1}p_{\mathbf{n}}+p_{\mathbf{n}}+1}^{2(j_{N-1}+1)p_{\mathbf{n}}} \quad \sum_{i_N=2j_Np_{\mathbf{n}}+1}^{2j_Np_{\mathbf{n}}+p_{\mathbf{n}}} \Delta_{\mathbf{i}}(x),$$

$$U(4,\mathbf{n},x,\mathbf{j}) = \sum_{\substack{i_k=2j_kp_{\mathbf{n}}+1\\ k=1,\ldots,N-2}}^{2j_kp_{\mathbf{n}}} \quad \sum_{i_{N-1}=2j_{N-1}p_{\mathbf{n}}+p_{\mathbf{n}}+1}^{2(j_{N-1}+1)p_{\mathbf{n}}} \quad \sum_{i_N=2j_Np_{\mathbf{n}}+p_{\mathbf{n}}+1}^{2(j_N+1)p_{\mathbf{n}}} \Delta_{\mathbf{i}}(x),$$

and so on. Finally

$$U(2^{N-1},\mathbf{n},x,\mathbf{j}) = \sum_{\substack{i_k=2j_kp_{\mathbf{n}}+p_{\mathbf{n}}+1\\ k=1,\ldots,N-1}}^{2(j_k+1)p_{\mathbf{n}}} \quad \sum_{i_N=2j_Np_{\mathbf{n}}+1}^{2j_Np_{\mathbf{n}}+p_{\mathbf{n}}} \Delta_{\mathbf{i}}(x),$$

(6)
$$U(2^N,\mathbf{n},x,\mathbf{j}) = \sum_{\substack{i_k=2j_kp_{\mathbf{n}}+p_{\mathbf{n}}+1\\ k=1,\ldots,N}}^{2(j_k+1)p_{\mathbf{n}}} \Delta_{\mathbf{i}}(x).$$

This blocking scheme is similar to that used in Tran (1990).
Now, for $\mathcal{J} = \{0,...,r_1-1\} \times ... \times \{0,...,r_N-1\}$ where $r_i = 2n_ip_{\mathbf{n}}^{-1}$, $i = 1,\ldots,N$, we define

(7)
$$T(\mathbf{n},x,i) = \sum_{\mathbf{j}\in\mathcal{J}} U(i,\mathbf{n},x,\mathbf{j})$$

and we write,

$$\widehat{f}_D^x - \mathrm{E}[\widehat{f}_D^x] = \frac{1}{\widehat{\mathbf{n}}\mathrm{E}[K_{\mathbf{1}}(x)]}\sum_{i=1}^{2^N} T(\mathbf{n}, x, i).$$

Note that, as raised by Biau and Cadre (2004), if one does not have the equalities $n_{\mathbf{i}} = 2r_{\mathbf{i}}p_{\mathbf{n}}$, the term say $T(\mathbf{n}, x, 2^N+1)$ (which contains the $\Delta_{\mathbf{i}}(x)$'s at the ends not included in the blocks above) can be added. This will not change the proof a lot.
Regarding (2.5), we can say that, for all $\eta > 0$

$$\mathbb{P}\left(|\widehat{f}_D^x - \mathrm{E}[\widehat{f}_D^x]| \geq \eta\right) \leq 2^N \max_{i=1,\dots,2^N} \mathbb{P}\left(T(\mathbf{n}, x, i) \geq \eta\widehat{\mathbf{n}}\mathrm{E}\left[K_{\mathbf{1}}(x)\right]\right).$$

Finally, the desired result follows from the evaluation of the following quantities :

$$\mathbb{P}\left(T(\mathbf{n}, x, i) \geq \eta\widehat{\mathbf{n}}\mathrm{E}\left[K_{\mathbf{1}}(x)\right]\right), \qquad \text{for all } i = 1, \dots, 2^N.$$

Without loss of generality, we will only consider the case $i = 1$. For this case, we enumerate the $M = \prod_{k=1}^N r_k = 2^{-N}\widehat{\mathbf{n}}p_{\mathbf{n}}^{-N} \leq \widehat{\mathbf{n}}p_{\mathbf{n}}^{-N}$ random variables $U(1, \mathbf{n}, x, \mathbf{j})$; $\mathbf{j} \in \mathcal{J}$, in an arbitrary way as $Z_1, \dots, Z_M$. Thus, for each Z_j it exists a certain $\mathbf{j}$ in $\mathcal{J}$ such that

$$Z_j = \sum_{\mathbf{i} \in I(1,\mathbf{n},x,\mathbf{j})} \Delta_{\mathbf{i}}(x),$$

where $I(1, \mathbf{n}, x, \mathbf{j}) = \{\mathbf{i} : 2j_k p_{\mathbf{n}} + 1 \leq i_k \leq 2j_k p_{\mathbf{n}} + p_{\mathbf{n}}; \quad k = 1, \dots, N\}$. Clearly the sets $I(1, \mathbf{n}, x, \mathbf{j})$ contain $p_{\mathbf{n}}^N$ sites and are separated by a distance of at least $p_{\mathbf{n}}$. So, according to Lemma 4.5.1 one can find independent random variables $Z_1^*, \dots Z_M^*$ having the same law as $(Z_j)_{j=1,\dots,M}$, such that

$$\sum_{j=1}^M \mathrm{E}|Z_j - Z_j^*| \leq 2CMp_{\mathbf{n}}^N\psi((M-1)p_{\mathbf{n}}^N, p_{\mathbf{n}}^N)\varphi(p_{\mathbf{n}}).$$

Therefore, by the Bernstein and Markov inequalities we have :

$$\mathbb{P}\left(T(\mathbf{n}, x, 1) \geq \eta\widehat{\mathbf{n}}\mathrm{E}\left[K_{\mathbf{1}}(x)\right]\right) \leq B_1 + B_2,$$

where

$$B_1 = \mathbb{P}\left(\left|\sum_{j=1}^M Z_j^*\right| \geq \frac{M\eta\widehat{\mathbf{n}}\mathrm{E}\left[K_{\mathbf{1}}(x)\right]}{2M}\right) \leq 2\exp\left(-\frac{(\eta\widehat{\mathbf{n}}\mathrm{E}\left[K_{\mathbf{1}}(x)\right])^2}{MVar\left[Z_1^*\right] + Cp_{\mathbf{n}}^N\eta\widehat{\mathbf{n}}\mathrm{E}\left[K_{\mathbf{1}}(x)\right]}\right)$$

and

$$\begin{aligned} B_2 &= \mathbb{P}\left(\sum_{j=1}^M |Z_j - Z_j^*| \geq \frac{\eta\widehat{\mathbf{n}}\mathrm{E}\left[K_{\mathbf{1}}(x)\right]}{2}\right) \\ &\leq \frac{1}{\eta\widehat{\mathbf{n}}\mathrm{E}\left[K_{\mathbf{1}}(x)\right]}\sum_{j=1}^M \mathrm{E}|Z_j - Z_j^*| \\ &\leq 2Mp_{\mathbf{n}}^N(\eta\widehat{\mathbf{n}}\mathrm{E}\left[K_{\mathbf{1}}(x)\right])^{-1}\psi((M-1)p_{\mathbf{n}}^N, p_{\mathbf{n}}^N)\varphi(p_{\mathbf{n}}). \end{aligned}$$

Since $\widehat{\mathbf{n}} = 2^N M p_{\mathbf{n}}^N$ and $\psi((M-1)p_{\mathbf{n}}^N, p_{\mathbf{n}}^N) \leq p_{\mathbf{n}}^N$ we get for $\eta = \eta_0 \sqrt{\dfrac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\, \phi_x(h_K)}}$

$$B_2 \leq \widehat{\mathbf{n}} p_{\mathbf{n}}^N \left(\log \widehat{\mathbf{n}}\right)^{-1/2} \left(\widehat{\mathbf{n}} \phi_x(h_K)\right)^{-1/2} \varphi(p_{\mathbf{n}}).$$

With $p_{\mathbf{n}} = C \left(\dfrac{\widehat{\mathbf{n}} \phi_x(h_K)}{\log \widehat{\mathbf{n}}} \right)^{1/2N}$, we can write

$$B_2 \leq \widehat{\mathbf{n}}\, \varphi(p_{\mathbf{n}}). \tag{8}$$

Consequently, from (H7), we have

$$\sum_{\mathbf{n}} \widehat{\mathbf{n}}\, \varphi(p_{\mathbf{n}}) < \infty.$$

Let us focus now on B_1. For this, let us evaluate asymptotically $Var\,[Z_1^*]$. Indeed,

$$Var\,[Z_1^*] = Var \left[\sum_{\mathbf{i} \in I(1,\mathbf{n},x,\mathbf{1})} \Delta_{\mathbf{i}}(x) \right] = \sum_{\mathbf{i},\mathbf{j} \in I(1,\mathbf{n},x,\mathbf{1})} Cov(\Delta_{\mathbf{i}}(x), \Delta_{\mathbf{j}}(x)).$$

Let $Q_{\mathbf{n}} = \sum_{\mathbf{i} \in I(1,\mathbf{n},x,\mathbf{1})} Var\,[\Delta_{\mathbf{i}}(x)]$ and $R_{\mathbf{n}} = \sum_{\mathbf{i} \neq \mathbf{j} \in I(1,\mathbf{n},x,\mathbf{1})} Cov(\Delta_{\mathbf{i}}(x), \Delta_{\mathbf{j}}(x))$. By assumptions (H1) and (H2), we have

$$Var[\Delta_{\mathbf{i}}(x)] \leq C(\phi_x(h_K) + (\phi_x(h_K))^2),$$

therefore

$$Q_{\mathbf{n}} = O\left(p_{\mathbf{n}}^N \phi_x(h_K)\right).$$

Concerning $R_{\mathbf{n}}$, we introduce the following sets :

$$S_1 = \{\mathbf{i},\mathbf{j} \in I(1,\mathbf{n},x,\mathbf{1}) : 0 < \|\mathbf{i}-\mathbf{j}\| \leq c_{\mathbf{n}}\},\ S_2 = \{\mathbf{i},\mathbf{j} \in I(1,\mathbf{n},x,\mathbf{1}) : \|\mathbf{i}-\mathbf{j}\| > c_{\mathbf{n}}\},$$

where $c_{\mathbf{n}}$ is a real sequence that converges to $+\infty$ and will be made precise later. Split $R_{\mathbf{n}}$ into two separate summations over sites in S_1 and S_2 :

$$\begin{aligned} R_{\mathbf{n}} &= \sum_{(\mathbf{i},\mathbf{j}) \in S_1} Cov\left(\Delta_{\mathbf{i}}(x), \Delta_{\mathbf{j}}(x)\right) + \sum_{(\mathbf{i},\mathbf{j}) \in S_2} Cov\left(\Delta_{\mathbf{i}}(x), \Delta_{\mathbf{j}}(x)\right) \\ &= R_{\mathbf{n}}^1 + R_{\mathbf{n}}^2. \end{aligned}$$

On one hand, we have :

$$\begin{aligned} R_{\mathbf{n}}^1 &\leq \sum_{(\mathbf{i},\mathbf{j}) \in S_1} \left|\mathrm{E}\left[K_{\mathbf{i}}(x) K_{\mathbf{j}}(x)\right] - \mathrm{E}\left[K_{\mathbf{i}}(x)\right] \mathrm{E}\left[K_{\mathbf{j}}(x)\right]\right| \\ &\leq C p_{\mathbf{n}}^N c_{\mathbf{n}}^N \phi_x(h_K) \Big((\phi_x(h_K))^{1/a} + \phi_x(h_K) \Big) \\ &\leq C p_{\mathbf{n}}^N c_{\mathbf{n}}^N \phi_x(h_K)^{(a+1)/a}. \end{aligned}$$

On the other hand, as the random variables $K_{\mathbf{j}}$ are bounded, we deduce from Lemma 4.5.2(ii) that

$$|Cov\left(\Delta_{\mathbf{i}}(x), \Delta_{\mathbf{j}}(x)\right)| \leq C\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right),$$

then

$$\begin{aligned} R_{\mathbf{n}}^2 &\leq C \sum_{(\mathbf{i},\mathbf{j})\in S_2} \varphi\left(\|\mathbf{i}-\mathbf{j}\|\right) \leq C p_{\mathbf{n}}^N \sum_{\mathbf{i}:\|\mathbf{i}\|\geq c_{\mathbf{n}}} \varphi\left(\|\mathbf{i}\|\right) \\ &\leq C p_{\mathbf{n}}^N c_{\mathbf{n}}^{-Na} \sum_{\mathbf{i}:\|\mathbf{i}\|\geq c_{\mathbf{n}}} \|\mathbf{i}\|^{Na} \varphi\left(\|\mathbf{i}\|\right). \end{aligned}$$

Let $c_{\mathbf{n}} = \left(\phi_x(h_K)\right)^{-1/Na}$, then we have

$$\begin{aligned} R_{\mathbf{n}}^2 &\leq C p_{\mathbf{n}}^N c_{\mathbf{n}}^{-Na} \sum_{\mathbf{i}:\|\mathbf{i}\|\geq c_{\mathbf{n}}} \|\mathbf{i}\|^{Na} \varphi\left(\|\mathbf{i}\|\right) \\ &\leq C p_{\mathbf{n}}^N \phi_x(h_K) \sum_{\mathbf{i}:\|\mathbf{i}\|\geq c_{\mathbf{n}}} \|\mathbf{i}\|^{Na} \varphi\left(\|\mathbf{i}\|\right). \end{aligned}$$

Because of (6) and (H2), we get $R_{\mathbf{n}}^2 \leq C p_{\mathbf{n}}^N \phi_x(h_K)$. Furthermore, under this choice of $c_{\mathbf{n}}$ we have $R_{\mathbf{n}}^1 \leq C p_{\mathbf{n}}^N \phi_x(h_K)$. Hence

$$Var\left[Z_1^*\right] = O\left(p_{\mathbf{n}}^N \phi_x(h_K)\right).$$

By using this last result, together with the definitions of $p_{\mathbf{n}}$, M and η, we get

$$B_1 \leq \exp\left(-C(\eta_0)\log\widehat{\mathbf{n}}\right).$$

Consequently, an appropriate choice of η_0 completes the proof of the first part of this lemma. Concerning the second part, it follows from $\mathrm{E}\left[\widehat{f_D^x}\right] = 1$, that

$$\mathrm{I\!P}\left(\left|\widehat{f_D^x}\right| \leq 1/2\right) \leq \mathrm{I\!P}\left(\left|\widehat{f_D^x} - \mathrm{E}\left[\widehat{f_D^x}\right]\right| \geq 1/2\right).$$

The result is then a consequence of the first part of the Lemma . ∎

Proof of Lemma 4.4.7. By the usual change of variables $u = \dfrac{y-t}{h_H}$, we have

$$\begin{aligned} \left|f^x(y) - \mathrm{E}\widehat{f_N^x}(y)\right| &= \left|\frac{1}{h_H \mathrm{E}K_{\mathbf{1}}(x)} \mathrm{E}\left(K\left(h_K^{-1}d(x,X_{\mathbf{1}})\right) E(H\left(h_H^{-1}(y-Y_{\mathbf{1}})\right)|X_{\mathbf{1}})\right) - f^x(y)\right| \\ &= \left|\frac{1}{h_H \mathrm{E}K_{\mathbf{1}}(x)} \mathrm{E}\left(K\left(h_K^{-1}d(x,X_{\mathbf{1}})\right) \int H\left(h_H^{-1}(y-t)\right) f^{X_{\mathbf{1}}}(t)dt\right) - f^x(y)\right| \\ &= \left|\frac{1}{\mathrm{E}K_{\mathbf{1}}(x)} \mathrm{E}\left(K\left(h_K^{-1}d(x,X_{\mathbf{1}})\right) \int H(u)\left(f^{X_{\mathbf{1}}}(y-uh_H) - f^x(y)\right)du\right)\right|. \end{aligned}$$

Hypotheses (H3) and (H6) allow to get

$$\left|f^x(y) - \mathrm{E}\widehat{f}_N^x(y)\right| \leq C(h_K^{b_1} + h_H^{b_2}),$$

this last yields the result. ∎

Proof of Lemma 4.4.8.

Using the fact that $[\alpha_x, \beta_x] \subset \bigcup_{k=1}^{\kappa_{\mathbf{n}}}(t_j - l_{\mathbf{n}}, t_j + l_{\mathbf{n}})$ with $l_{\mathbf{n}} = \widehat{\mathbf{n}}^{-\frac{3}{2}\alpha-\frac{1}{2}}$ and $\kappa_{\mathbf{n}} \leq \widehat{\mathbf{n}}^{\frac{3}{2}\alpha+\frac{1}{2}}$. Set, $j(y) = \arg\min_{j\in\{1,2,\ldots,\kappa_{\mathbf{n}}\}} |y - t_j|$ and consider the following decomposition

$$\left|\widehat{f}_N^x(y) - \mathrm{E}\widehat{f}_N^x(y)\right| \leq \underbrace{\left|\widehat{f}_N^x(y) - \widehat{f}_N^x(t_{j(y)})\right|}_{T_1} + \underbrace{\left|\widehat{f}_N^x(t_{j(y)}) - \mathrm{E}\widehat{f}_N^x(t_{j(y)})\right|}_{T_2} + \underbrace{\left|\mathrm{E}\widehat{f}_N^x(t_{j(y)}) - \mathrm{E}\widehat{f}_N^x(y)\right|}_{T_3}.$$

- Concerning (T_1) and (T_3), we use the fact that H is bounded :

$$\begin{aligned}\left|\widehat{f}_N^x(y) - \widehat{f}_N^x(t_{j(y)})\right| &\leq \frac{h_H^{-1}}{\widehat{\mathbf{n}}\mathrm{E}\left[K_{\mathbf{1}}(x)\right]} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K_{\mathbf{i}}(x)\left|H_{\mathbf{i}}(y) - H_{\mathbf{i}}(t_{j(y)})\right| \\ &\leq C\frac{h_H^{-1}}{\widehat{\mathbf{n}}\mathrm{E}\left[K_{\mathbf{1}}(x)\right]} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K_{\mathbf{i}}(x)\left|H_{\mathbf{i}}(y) - H_{\mathbf{i}}(t_{j(y)})\right| \\ &\leq C\frac{l_{\mathbf{n}}}{h_H^2}\widehat{f}_D^x \leq \frac{Cl_{\mathbf{n}}}{h_H^2}.\end{aligned}$$

Notice that, this last inequality comes from the almost complete consistency of $\widehat{f}_D^x$ given in Lemma 4.4.6. Using the definition of $l_{\mathbf{n}}$ and (H7), one can write $\dfrac{l_{\mathbf{n}}}{h_H^2} = o\left(\sqrt{\dfrac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H\phi_x(h_K)}}\right)$. Hence

$$\left|\widehat{f}_N^x(y) - \widehat{f}_N^x(t_{j(y)})\right| = o\left(\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H\phi_x(h_K)}}\right), \qquad a.co. \tag{9}$$

and

$$\left|\mathrm{E}\widehat{f}_N^x(t_{j(y)}) - \mathrm{E}\widehat{f}_N^x(y)\right| = o\left(\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H\phi_x(h_K)}}\right). \tag{10}$$

- Looking now at the term (T_2), we can write, $\forall\ \eta > 0$

$$\begin{aligned}
&\mathbb{P}\left(\sup_{y\in(\alpha_x,\beta_x)}\left|\widehat{f}_N^x(t_{j(y)})-\mathrm{E}\widehat{f}_N^x(t_{j(y)})\right|>\eta\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H\phi_x(h_K)}}\right)\\
&\quad=\mathbb{P}\left(\max_{j\in\{1,2,\ldots,\kappa_{\mathbf{n}}\}}\left|\widehat{f}_N^x(t_j)-\mathrm{E}\widehat{f}_N^x(t_j)\right|>\eta\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H\phi_x(h_K)}}\right)\\
&\quad\leq\kappa_{\mathbf{n}}\max_{j\in\{1,2,\ldots,\kappa_{\mathbf{n}}\}}\mathbb{P}\left(\left|\widehat{f}_N^x(t_j)-\mathrm{E}\widehat{f}_N^x(t_j)\right|>\eta\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H\phi_x(h_K)}}\right).
\end{aligned}$$

Setting $\Delta_{\mathbf{i}}' = K_{\mathbf{i}}(x)H_{\mathbf{i}}(t_j) - \mathrm{E}\left[K_{\mathbf{i}}(x)H_{\mathbf{i}}(t_j)\right]$.
Thus, the claimed result in T_2 follows from, a straightforward modification of the spatial decomposition (7). Indeed, we consider the same decomposition with $p_{\mathbf{n}} = C\left(\frac{\widehat{\mathbf{n}}h_H\phi_x(h_K)}{\log\widehat{\mathbf{n}}}\right)^{1/2N}$.
So, by using the fact that $Var\left[\Delta_{\mathbf{i}}'\right] = O(h_H\phi_x(h_K))$, we get

$$\sum_{\mathbf{i}\in I(1,\mathbf{n},x,\mathbf{1})} Var\left[\Delta_{\mathbf{i}}'(x)\right] = O\left(p_{\mathbf{n}}^N h_H\phi_x(h_K)\right).$$

While, the use of the fact that $\mathrm{E}\left[H_{\mathbf{i}}H_{\mathbf{j}}|(X_{\mathbf{i}},X_{\mathbf{j}})\right] = O(h_H^2)\ \forall;\mathbf{i}\neq\mathbf{j}$. allows us to write as in the proof of Lemma 4.4.6, by taking a sequence $c_{\mathbf{n}}\to+\infty$:

$$\sum_{\mathbf{i}\neq\mathbf{j}\in I(1,\mathbf{n},x,\mathbf{1})}\left|Cov(\Delta_{\mathbf{i}}'(x),\Delta_{\mathbf{j}}'(x))\right| \leq Cp_{\mathbf{n}}^N\left(c_{\mathbf{n}}^N h_H^2\phi_x(h_K)^{1/a}\phi_x(h_K) + c_{\mathbf{n}}^{-Na}\sum_{\mathbf{i}:\|\mathbf{i}\|\geq c_{\mathbf{n}}}\|\mathbf{i}\|^{Na}\varphi\left(\|\mathbf{i}\|\right)\right).$$

Now, choosing $c_{\mathbf{n}} = (h_H\phi_x(h_K))^{-1/Na}$ permits to get

$$\sum_{\mathbf{i}\neq\mathbf{j}\in I(1,\mathbf{n},x,\mathbf{1})}\left|Cov(\Delta_{\mathbf{i}}'(x),\Delta_{\mathbf{j}}'(x))\right| \leq Cp_{\mathbf{n}}^N h_H\phi_x(h_K).$$

Thus,

$$Var\left[\sum_{\mathbf{i}\in I(1,\mathbf{n},x,\mathbf{1})}\Delta_{\mathbf{i}}'(x)\right] = O\left(p_{\mathbf{n}}^N h_H\phi_x(h_K)\right).$$

Finally, the rest of the proof follows the same lines as the proof of Lemma 4.4.6 when we apply Lemma 4.5.1, Bernstein and Markov inequalities which permit to write

$$\sup_{y\in(\alpha_x,\beta_x)}\left|\widehat{f}_N^x(t_{j(y)})-\mathrm{E}\widehat{f}_N^x(t_{j(y)})\right| = O_{a.co.}\left(\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H\phi_x(h_K)}}\right).$$

Finally, the Lemma is a consequence of this last result, (9) and (10). ■

Acknowledgements. The authors thank the Associate Editor and the anonymous referees immensely for a careful reading of the paper.

Bibliographie

Biau, G. and Cadre, B. (2004). Nonparametric Spatial Prediction. *Statist. Inf. Stoch. Proc*, **7**, 327-349.
Bosq, D. (2000). *Linear Processes in Function Spaces : Theory and applications.* Lecture Notes in Statistics, 149, Springer.
Carbon,M., Tran, L.T. and Wu, B. (1997). Kernel density estimation for random fields : the L_1 theory. *J. Nonparametric. Statist,* **6**, 157-170.
Collomb, G., Härdle, W. and Hassani, S. (1987). A note on prediction via conditional mode estimation. *J. Statist. Plann. Inference*, **15**, 227-236.
Cressie, N.A.C. (1991). *Statistics for Spatial Data.* Wiley Series in Probability and Mathematical Statistics, New-York.
Dabo-Niang, S. and Yao, A-F. (2007). Kernel regression estimation for continuous spatial processes. *Math. Methods. Statist.* **16**, 298-317.
Dabo-Niang, S. and Thiam, B. (2010). L_1 consistency of a kernel estimate of spatial conditional quantile. *Statist. Probab. Lett*, **80**, 1447-1458.
Dabo-Niang, S., Rachdi, M. and Yao, A-F. (2011a). Spatial kernel regression estimation and prediction for functional random fields. *Far. East. J. Statist*, **37**, No. 2, 77-113.
Dabo-Niang, S., Kaid, Z. and Laksaci A. (2011b). Spatial conditional quantile regression : Weak consistency of a kernel estimate. Preprint.
Dabo-Niang, S. and Laksaci, A. (2010). Note on conditional mode estimation for functional dependent data, *Statistica*, **70**, 83-94.
Ezzahrioui, M. and Ould-Said, E. (2008). Asymptotic normality of nonparametric estimators of the conditional mode function for functional data *J. Nonparametric. Statist*, **20**, 3-18.
Ferraty, F., Laksaci, A. and Vieu, P.(2005). Functional times series prediction via conditional mode. *C. R., Math., Acad. Sci. Paris*, **340**, No. 5, 389-392.
Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis.* Springer Series in Statistics, Springer New York.
Ferraty, F., Laksaci, A., Tadj, A. and Vieu, P. (2010). Rate of uniform consistency for nonparametric estimates with functional variables. *J. Statist. Plann. Inference,* **140**, 335-352.
Guyon, X. (1987). Estimation d'un champ par pseudo-vraisemblance conditionnelle :

Etude asymptotique et application au cas Markovien. *Proceedings of the Sixth Franco-Belgian Meeting of Statisticia.*
HAINING, R. (2003) Spatial Data Analysis : Theory and Practice. Combridge University Press
LAKSACI, A. and MAREF, F. (2009). Estimation non paramï£¡trique de quantiles conditionnels pour des variables fonctionnelles spatialement dï£¡pendantes, *C. R., Math., Acad. Sci. Paris,* **347**, 1075-1080.
LI, J. and TRAN, L.T. (2009). Nonparametric estimation of conditional expectation. *J. Statist. Plann. Inference,* **139**, 164-175.
RAMSAY, J.O. and SILVERMAN, B.W. (2005). *Functional Data Analysis.* Second Edition. Springer, New York.
RAMSAY, J. (2008). Fda problems that I like to talk about. Personal communication.
SARDA, P. and VIEU, P. (2000). Kernel regression. In M. Schimek (ed.) *Smoothing and regression ; Approaches, Computation, and Application.*, Wiley Series in Probability and Statistics, Wiley, New-York.
TRAN, L.T. (1990). Kernel density estimation on random fields, *J. Multivariate Anal*, **34**, 37-53.

Chapitre 3

La normalité asymptotique et la convergence en norme L^p de l'estimateur spatial du mode conditionnel

Dans ce chapitre nous nous intéressons à la convergence en norme L^p et la normalité asymptotique de la version spatiale de l'estimateur à noyau du mode conditionnel avec une variable explicative fonctionnelle. De même que le chapitre précédent, la difficulté technique réside dans l'absence d'une mesure de référence classique. De plus, en statistique non-paramétrique vectorielle, l'étude de la normalité asymptotique nécessite un calcul asymptotique exact des termes du biais et de la variance qui est basée sur la dérivabilité des modèles non-paramétriques. Cependant, la structure topologique de notre espace fonctionnel est trop faible pour définir cette notion. En effet, la condition de dérivabilité exige une structure topologique de type banachique, or ici nous considérons un espace fonctionnel semi-métrique. Ainsi, l'impact principal de cette contribution est la généralisation en dimension infinie des résultats déjà existants en statistique multi-variée mais sous des conditions moins restrictives. *Ce chapitre fait l'objet d'un article soumis pour publication.*

Asymptotic properties of the kernel estimate of the spatial conditional mode when the regressor is functional

Sophie Dabo-Niang ♯*, Zoulikha Kaid ♯♮
Labo. EQUIPPE, Maison de la recherche
♯Université Lille 3,
BP 60149, 59653,Villeneuve d'Ascq cedex, France.
e-mail: sophie.dabo@univ-lille3.fr, zoulikha.kaid@etu.univ-lille3.fr

Ali Laksaci ♮
♮ Agence Nationale de Dveloppement
et de Recherche Universitaire (A.N.D.R.U.)
Laboratoire de Mathématiques
Univ. Djillali Liabès
BP 89, 22000 Sidi Bel Abbès, Algeria
e-mail: alilak@yahoo.fr

August 23, 2012

Abstract
The kernel method is used to propose an estimator of the spatial modal regression for functional regressors, considering a mixing spatial structure for $\{(X_{\mathrm{i}}, Y_{\mathrm{i}}),\ \mathrm{i} \in \mathbb{Z}^N\}$, $N \geq 1$. We establish, under some general mixing conditions, the L^p-consistency and the asymptotic normality of the estimator. The performance of the proposed estimator for finite samples is checked in a simulation study.

Keywords: Spatial process, conditional mode estimate, non-parametric, Functional data.

*corresponding author

3.1 Introduction

Let $(Z_{\mathbf{i}} = (X_{\mathbf{i}}, Y_{\mathbf{i}}),\, \mathbf{i} \in \mathbb{Z}^N)$ be a $\mathcal{F} \times \mathbb{R}$-valued measurable strictly stationary spatial process, defined on a probability space $(\Omega,\ \mathcal{A}, \mathbb{P})$, where $\mathcal{F}$ is a semi-metric space. Let d denote the semi-metric and $N \geq 1$. We assume that the $Z_{\mathbf{i}}$'s have the same distribution as (X, Y) and there exists a regular version of the conditional probability of Y given X and its absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$, has bounded density. Moreover, we suppose that for a given x the conditional density f^x of Y given $X = x$ is unimodal in some fixed compact S and the conditional mode, denoted by $\theta(x)$ is defined by

$$\theta(x) = \arg\max_{y \in S} f^x(y).$$

Now, assuming that the process under study $(Z_{\mathbf{i}})_{\mathbf{i}}$ is observed over a rectangular domain $I_{\mathbf{n}} = \left\{\mathbf{i} = (i_1, ..., i_N) \in \mathbb{Z}^N\ ,\ \ 1 \leq i_k \leq n_k,\ k = 1, ..., N\right\}$, $\mathbf{n} = (n_1, ..., n_N) \in \mathbb{Z}^N$, a natural and usual estimator of $\theta(x)$ denoted $\widehat{\theta}(x)$, is given by :

$$\widehat{\theta}(x) \ = \ \arg\sup_{y \in S} \widehat{f}^x(y), \tag{1}$$

where $\widehat{f}^x(\cdot)$ is the estimator of $f^x(\cdot)$ defined by

$$\widehat{f}^x(y) = \frac{h_H^{-1} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K(h_K^{-1} d(x, X_{\mathbf{i}})) H(h_H^{-1}(y - Y_{\mathbf{i}}))}{\sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K(h_K^{-1} d(x, X_{\mathbf{i}}))}, \qquad \forall y \in \mathbb{R}$$

with K and H are kernels functions and $h_K = h_{K,\mathbf{n}}$ (resp. $h_H = h_{H,\mathbf{n}}$) is a sequence of positive real numbers.
The main goal of this paper is to study the asymptotic proprieties of the nonparametric estimate $\widehat{\theta}(x)$ of $\theta(x)$ when the explanatory variable X is valued in the space $\mathcal{F}$ of eventually infinite dimension and the functional random filed $(Z_{\mathbf{i}}, \mathbf{i} \in \mathbb{Z}^N)$ satisfies the following mixing condition :

$$\left\{ \begin{array}{c} \text{There exists a function } \varphi\left(t\right) \downarrow 0 \text{as } t \to \infty, \text{such that} \\ \forall\, E,\ E^{'} \text{subsets of } \mathbb{Z}^N \text{with finite cardinals} \\ \alpha\left(\mathcal{B}\left(E\right), \mathcal{B}\left(E^{'}\right)\right) = \displaystyle\sup_{B \in \mathcal{B}(E),\, C \in \mathcal{B}\left(E^{'}\right)} \left|\mathbb{P}\left(B \cap C\right) - \mathbb{P}\left(B\right)\mathbb{P}\left(C\right)\right| \\ \leq \psi\left(\text{Card}\left(E\right), \text{Card}\left(E^{'}\right)\right) \varphi\left(\text{dist}\left(E, E^{'}\right)\right), \end{array} \right. \tag{2}$$

where $\mathcal{B}\left(E\right)$($resp.$ $\mathcal{B}\left(E^{'}\right)$) denotes the Borel σ-field generated by $(Z_{\mathbf{i}},\ \mathbf{i} \in E)$ ($resp.$ $\left(Z_{\mathbf{i}},\ \mathbf{i} \in E^{'}\right)$), Card$(E)$ ($resp.$ Card$\left(E^{'}\right)$) the cardinality of E ($resp.$ $E^{'}$), dist$\left(E, E^{'}\right)$ the Euclidean distance between E and $E^{'}$ and $\psi : \mathbb{Z}^2 \to \mathbb{R}^+$ is a symmetric positive function nondecreasing in each variable such that

$$\psi\left(n, m\right) \leq C \min\left(n, m\right), \qquad \forall n, m \in \mathbb{Z} \tag{3}$$

for some $C > 0$.
Noting that the statistical problems involved in the modelization of spatial data have received an increasing interest in the literature. Key references on spatial statistic are Ripley (1981), Cressie (1993), or Diggle and Ribeiro (2007). The nonparametric treatment of such data is relatively recent. The first results have been obtained by Tran (1990). For relevant works on the nonparametric modelization of spatial data, see Lu and Chen (2004), Biau and Cadre (2004), Carbon *et al.* (2007) or Tran (2009), Dabo-Niang and Thiam (2010).
Currently, the progress of informatics tools and the modern technology permits the recovery of increasingly bulky data which are recorded densely over time. They are typically treated as curve or functional data. This presents the advantage to give a framework which fits better to the functional nature of the observations. For an overview on functional data analysis, we refer the reader to the monographs of Ramsay and Silverman (2002 and 2005), Bosq (2000) for parametric models and Ferraty and Vieu (2006) for the nonparametric case. In this context, the spatial analysis has been selected by Ramsay (2008) among the eight most interesting research subjects in functional data analysis. This great consideration is motivated by the increasing number of situations coming from different fields of applied sciences for which the data are of functional nature and showing a spatial interaction.
Despite its importance in applications, the nonparametric functional spatial prediction has not yet been fully explored. Specifically, this problem has been investigated by Dabo-Niang *et al* . (2011a), they stated the weak and the strong convergence of the functional spatial kernel estimate of the regression function. While Laksaci and Maref (2009) obtained a rate of almost complete convergence of the kernel estimate of the spatial conditional quantiles where the regressor is of functional nature. The asymptotic normality of this estimate has been stated by Dabo-Niang *et al.* (2011b). Among the recent papers on the conditional mode estimation in functional statistics, we refer to Ferraty and Vieu (2006), Dabo-Niang and Laksaci (2010), Ezzahrioui and Ould-Said (2008) and Ferraty *et al.* (2010).
The main aim of this paper is to study under general conditions, the asymptotic proprieties of the functional spatial kernel estimate of the conditional mode function. More precisely, we prove the p-integrated consistency by giving the upper bounds for the estimation error. In addition, we establish the asymptotic normality of the estimator considered. We point out that, our asymptotic results are useful in some statistical problem such as the prediction problems and the determination of confidence intervals. The present work extended to spatial case the result of Dabo-Niang and Laksaci (2010) given in functional time series case. It should be noted that extending classical nonparametric statistic results for functional random fields is far from being trivial and the main difficulties that arise in the analysis of spatial data comes from the fact that points in the $N-$dimensional space do not have a linear order. Furthermore, the practice interest of our study comes mainly from the fact that the main fields of application of functional statistical methods are related to the analysis of continuously indexed spatial processes.
The paper is organized as follows : the next section is dedicated to the L_p convergence. Section 3 is devoted to the asymptotic normality results of the spatial conditional mode estimate. In Section 5, we discuss the impact of our asymptotic result in some statistical

problems such as the choice of the smoothing parameters, the determination of confidence intervals and the prediction problem. The performance of our approach for finite samples is also checked by a simulation study in this Section. The proofs of the auxiliary results are relegated to the Appendix.

3.2 Weak consistency

In this section, we focus on pointwise consistency in p-mean. To do that, we denote by C or C' some strictly positive generic constants, $g^{(j)}$ the j^{th} order derivative of the function g and we consider the following hypotheses :

(H1) $\forall r>0, \mathbb{P}(X\in B(x,r)) =: \phi_x(r)>0$. Moreover, $\phi_x(r)\longrightarrow 0$ as $r\longrightarrow 0$.

(H2) For all $k\geq 2$, we suppose that :

$$\begin{cases} \text{(i) For all } 1\leq \mathbf{i}_1<\dots\mathbf{i}_k\leq n, \text{ the conditional density of } (Y_{\mathbf{i}_1},\dots Y_{\mathbf{i}_k}) \text{ given } (X_{\mathbf{i}_1},\dots X_{\mathbf{i}_k}) \\ \qquad\qquad \text{exists and is uniformely bounded} \\ \text{(ii) There exists } v_k>0, \text{ such that} \\ \qquad \max\left(\max_{1\leq \mathbf{i}_1<\dots\mathbf{i}_k\leq n} P\left(d(X_{\mathbf{i}_j},x)\leq r, 1\leq j\leq k\right), \phi_x^k(r)\right) = O\left(\phi_x^{1+v_k}(r)\right). \end{cases}$$

(H3) The mixing coefficient $\varphi(\cdot)$ satisfying

$$\exists\delta>0, \quad \text{such that } \sum_{i=1}^{\infty} i^{\delta}\varphi(i)<\infty$$

(H4) f^x is 2-times continuously differentiable with respect y on $\mathbb{R}$ such that, $\forall(y_1,y_2)\in\mathbb{R}^2$, $\forall(x_1,x_2)\in N_x\times N_x$,

$$|f^{x_1\,(j)}(y_1)-f^{x_2\,(j)}(y_2)|\leq C\left(d(x_1,x_2)^{b_1}+|y_1-y_2|^{b_2}\right), \quad C>0,\ b_1>0,\ b_2>0 \text{ for } j=0,1,2$$

with the convention $f^{x\,(0)}=f^x$.

(H5) K is a function with support $(0,1)$ such that $0<C'<K(t)<C<\infty$.

(H6) H is of class $\mathcal{C}^2$, of compact support and satisfies

$$\int H(t)dt=1, \quad \text{and}$$
$$\forall(y_1,y_2)\in\mathbb{R}^2, \quad |H^{(j)}(y_1)-H^{(j)}(y_2)|\leq C|y_1-y_2|, \quad j=0,1,2 \text{ with } H^{(0)}=H.$$

(H7) There exists $0<\alpha<(\delta-5N)/3N$ and $\eta_0>0$, such that $\lim_{\mathbf{n}\to\infty}\widehat{\mathbf{n}}^{\alpha}\,h_H=\infty$ and $C\widehat{\mathbf{n}}^{\frac{(5+3\alpha)N-\delta}{\delta}+\eta_0}\leq h_H\phi_x(h_K)$

Theorem 1 *Under hypotheses (H1)-(H7), and if $nh_H^3\phi_x(h_K)\to\infty$, we have for all $p\in[1,\infty[$*

$$\left\|\widehat{\theta}(x)-\theta(x)\right\|_p = O\left(h_K^{b_1}+h_H^{b_2}\right)+O\left(\left(\frac{1}{\widehat{\mathbf{n}}\,h_H^3\phi_x(h_K)}\right)^{\frac{1}{2}}\right)$$

where $\|.\|_p=(E^{1/p}|.|^p)$.

Proof of Theorem 1. For all $\mathbf{i} \in \mathcal{I}_{\mathbf{n}}$, we put $K_{\mathbf{i}} = K\left(h_K^{-1} d(x, X_{\mathbf{i}})\right)$, $H_{\mathbf{i}}(y) = H\left(h_H^{-1}(y - Y_{\mathbf{i}})\right)$ and

$$\widehat{f_N^x}(y) = \frac{1}{\widehat{\mathbf{n}}\, h_H E K_{\mathbf{1}}} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K_{\mathbf{i}} H_{\mathbf{i}}(y), \qquad \widehat{f_D^x} = \frac{1}{\widehat{\mathbf{n}}\, E K_{\mathbf{1}}} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K_{\mathbf{i}}.$$

By a Taylor development of the function $\widehat{f_N^x}^{(1)}$ at $\theta(x)$ we write

$$\widehat{f_N^x}^{(1)}(\widehat{\theta}(x)) = \widehat{f_N^x}^{(1)}(\theta(x)) + (\widehat{\theta}(x) - \theta(x)) \widehat{f_N^x}^{(2)}(\theta^*(x))$$

where $\theta^*(x)$ is between $\widehat{\theta}(x)$ and $\theta(x)$. Observe that the unimodality of f^x and assumption (H6) allows to write that

$$f^{x\,(1)}(\theta(x)) = \widehat{f_N^x}^{(1)}(\widehat{\theta}(x)) = 0 \qquad \text{and} \qquad f^{x\,(2)}(\theta(x)) < 0.$$

Therefore, if $\widehat{f_D^x} \neq 0$, we have

$$\widehat{\theta}(x) - \theta(x) = \frac{1}{\widehat{f_N^x}^{(2)}(\theta^*(x))} \left(\widehat{f_N^x}^{(1)}(\theta(x)) - f^{x\,(1)}(\theta(x)) \right). \tag{4}$$

By following the same idea's as those used by Dabo-Niang et al.(2012) in Lemma 3.4 we show that

$$\sup_{y \in S} |\widehat{f_N^x}^{(2)}(y) - f^{x\ (2)}(y)| \to 0. \qquad a.co.$$

We combine this consistency with the fact that

$$\widehat{\theta}(x) - \theta(x) \to 0 \qquad \text{almost completely}$$

we obtain

$$\exists C > 0 \quad \text{such that} \quad \left| \frac{1}{\widehat{f_N^x}^{(2)}(\theta^*(x))} \right| \leq C \quad a.s.$$

It follows that

$$\left\| \widehat{\theta}(x) - \theta(x) \right\|_p \leq C \left\| \widehat{f_N^x}^{(1)}(\theta(x)) - f^{x\,(1)}(\theta(x)) \right\|_p + \left(P\left(\widehat{f_D^x} = 0 \right) \right)^{1/p}.$$

Theorem 1 is then a consequence of the following lemmas.

Lemma 3.2.1 *Under hypotheses (H1)-(H3), (H5)-(H6), we have,*

$$\left\| \widehat{f_N^x}^{(1)}(\theta(x)) - E\left[\widehat{f_N^x}^{(1)}(\theta(x)) \right] \right\|_p = O\left(\sqrt{\frac{1}{\widehat{\mathbf{n}} h_H^3 \phi_x(h_K)}} \right). \tag{5}$$

Lemma 3.2.2 *If the hypotheses (H1), (H4)-(H6) are satisfied, we get*

$$E\left[\widehat{f}_N^{x\,(1)}(\theta(x))\right] - f^{x\,(1)}(\theta(x)) = O(h_K^{b_1} + h_H^{b_2}).$$

Lemma 3.2.3 *Under the conditions of Lemma 3.2.1, we get*

$$\left(P\left(\widehat{f}_D^x = 0\right)\right)^{1/p} = O\left(\sqrt{\frac{1}{\widehat{\mathbf{n}}\phi_x(h_K)}}\right).$$

3.3 Asymptotic normality

In order to establish our asymptotic results we need the following additional assumptions :

(H8) The bandwidth h_K satisfies :

$$h_K \downarrow 0, \quad \forall t \in [0,1] \quad \lim_{h_K \to 0} \frac{\phi_x(th_K)}{\phi_x(h_K)} = \beta_x(t) \text{ and } \widehat{\mathbf{n}} h_H \phi_x(h_K) \to \infty \text{ as } \mathbf{n} \to \infty.$$

(H9) The functions $\Phi(s) = E\left[\frac{\partial^2 f^X(\theta(x))}{\partial y^2} - \frac{\partial^2 f^x(\theta(x))}{\partial y^2}\middle| d(x,X) = s\right]$ are derivable at $s = 0$.

(H10) The kernel K from $\mathbb{R}$ into $\mathbb{R}^+$ is a differentiable function supported on $[0,1]$. Its derivative K' exists and is such that there exist two constants C and C' with $-\infty < C < K'(t) < C' < 0$ for $0 \le t \le 1$.

(H11) There exists

$$\frac{N}{\delta} < \theta < 1, \text{such that} \quad \widehat{\mathbf{n}}^{(-1+\theta_1)/(1+2N)} \le h_H \phi_x(h_K).$$

Theorem 2 *Assume that (H1), (H3)-(H11) hold, then we have for any $x \in \mathcal{A}$,*

$$\left(\frac{\widehat{\mathbf{n}} h_H^3 \phi_x(h_K)}{\sigma_1^2(x)}\right)^{1/2} \left(\widehat{\theta}(x) - \theta(x) - B_{\mathbf{n}}(x)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad as \quad \mathbf{n} \to \infty.$$

where

$$B_{\mathbf{n}}(x) = B_H(x) h_H + B_K(x) h_K$$

with

$$\begin{aligned} B_H(x) &= \int t H(t) dt \\ B_K(x) &= \Phi_0'(0) \frac{\left(K(1) - \int_0^1 (sK(s))' \beta_x(s) ds\right)}{f^{x(2)}(\theta(x)) \left(K(1) - \int_0^1 K'(s) \beta_x(s) ds\right)}. \end{aligned}$$

and

$$\sigma_1^2(x) = \frac{\beta_2 f^x(\theta(x))}{\beta_1^2 (f^{x(2)}(\theta(x)))^2} \int H^2(t) dt \ \ (\textit{with } \beta_j = K^j(1) - \int_0^1 (K^j)'(s) \beta_x(s) ds, \ \textit{for}, \ j = 1, 2),$$

where

$$\mathcal{A} = \{x \in \mathcal{F},\ f^{x(2)}(\theta(x))f^{x}(\theta(x)) \neq 0\}$$

and $\xrightarrow{\mathcal{D}}$ *means the convergence in distribution.*

Obviously, if one imposes some additional assumptions on the function $\phi_x(\cdot)$ and the bandwidth parameters (h_K and h_H) we can improved our asymptotic normality by removing the bias term $B_{\mathbf{n}}(x)$.

Corollary 1 *Under the hypotheses of Theorem 2 and if the bandwidth parameters (h_K and h_H) and the function $\phi_x(h_K)$ satisfy :*

$$\lim_{\mathbf{n}\to\infty}(h_H + h_K)\sqrt{\widehat{\mathbf{n}}h_H^3\phi_x(h_K)} = 0$$

we have

$$\left(\frac{\widehat{\mathbf{n}}h_H^3\phi_x(h_K)}{\sigma_1^2(x)}\right)^{1/2}\left(\widehat{\theta}(x) - \theta(x)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad as \quad \mathbf{n}\to\infty.$$

Proof of Theorem and Corollary By using the same analytical arguments as those used in the previous we show that Theorem 2 and Corollary 2 are a consequence of Lemma 4.4.6, remark (1) and the following results.

Lemma 3.3.1 *Under the hypotheses of Theorem 2, we have*

$$E\left[\widehat{f}_N^{x\,(1)}(\theta(x))\right] = f^{x(2)}(\theta(x))(B_H h_H + B_K h_K) + o(h_H) + o(h_K)$$

Remark 1

Observe that, the result of this lemma permits to write

$$\mathrm{E}\left[\widehat{f}_N^{x\,(1)}(\theta(x))\right] = O(h_H) + O(h_K)$$

Lemma 3.3.2 *Under the hypotheses of Theorem 2, we have*

$$Var\left[\widehat{f}_N^{x\,(1)}(\theta(x))\right] = \frac{(f^{x(2)}(\theta(x)))^2\sigma_1^2(x)}{\widehat{\mathbf{n}}h_H^3\phi_x(h_K)} + o\left(\frac{1}{\widehat{\mathbf{n}}h_H\phi_x(h_K)}\right),$$

Lemma 3.3.3 *Under the hypotheses of Theorem 2*

$$\left(\frac{\widehat{\mathbf{n}}h_H^3\phi_x(h_K)}{\sigma^2(x)}\right)^{1/2}\left(\widehat{f}_N^{x\,(1)}(\theta(x)) - E\left[\widehat{f}_N^{x\,(1)}(\theta(x))\right]\right) \to N(0,1)$$

where $\sigma^2(x) = f^{x(2)}(\theta(x)))^2\sigma_1^2(x)$

3.4 Discussion and applications

3.4.1 Discussion

The purpose of this section is to show the applicability of our model in practice. Firstly, we discuss the practical determination of the different parameters intervening in the computation of the estimator, such as the spatial dependency and the smoothing parameters choice. The finite sample performance of our approach is checked in the last part of this paragraph.
How to control the spatial correlation in practice? Often, the spatial dependence is modeled through the covariogram or the variogram function (see Cressie 1993). Specifically, in functional statistic we can apply another function so-called trace-variogram function (see Giraldo 2009). As suggested by Dabo-Ninang *et al.* (2011a) the appropriate manner to control the spatial correlation in practice is to integrate it in the computation of the estimator by taking $\theta(X_{\mathbf{k}})$, for a fixed site $\mathbf{k}$, as

$$\widehat{\theta}(X_{\mathbf{k}}) = \arg\min \frac{h_H^{-1} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K(h_K^{-1} d(X_{\mathbf{k}}, X_{\mathbf{i}})) H(h_H^{-1}(y - Y_{\mathbf{i}})) \mathbb{1}_{W_{\mathbf{k}}}(\mathbf{i})}{\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K(h_K^{-1} d(X_{\mathbf{k}}, X_{\mathbf{i}})) \mathbb{1}_{W_{\mathbf{k}}}(\mathbf{i})}$$

where $W_{\mathbf{k}}$ is a vicinity set defined by

$$\text{(6)} \qquad \text{for all } \mathbf{k} \quad W_{\mathbf{k}} = \{\mathbf{i}, \text{ such that } \gamma(\mathbf{i}, \mathbf{k}) \le \iota_{\mathbf{n}}\}; \quad \gamma \quad \text{is the trace-varigram function.}$$

Noting that this methodology require the estimation of γ, which can be obtained empirically by

$$\widehat{\gamma}(\mathbf{l}, \mathbf{k}) = \frac{1}{2\#N_{\mathbf{l},\mathbf{k}}} \sum_{\mathbf{i},\mathbf{j}\in N_{\mathbf{l},\mathbf{k}}} d(X_{\mathbf{i}}, X_{\mathbf{j}})$$

where $N_{\mathbf{l},\mathbf{k}} = \{\mathbf{i}, \mathbf{j} \in \mathcal{I}_{\mathbf{n}} \text{ such that } \|\mathbf{i} - \mathbf{j}\| = \|\mathbf{l} - \mathbf{k}\|\}$ and $\#N_{\mathbf{l},\mathbf{k}}$ is the cardinal of $N_{\mathbf{l},\mathbf{k}}$. It should be noted that the advantage of this approach is that the functional nature of the data is also taken into account in the quantification of the spatial correlation. Alternatively, in the isotropic case where, the dependence is just a function of the distance between the locations, we can proceed with the following vicinity set

$$\text{(7)} \qquad \text{for all } \mathbf{k} \quad V_{\mathbf{k}} = \{\mathbf{i}, \text{ such that } \|\mathbf{i} - \mathbf{k}\| \le \nu_{\mathbf{n}}\}$$

where $\nu_{\mathbf{n}}$ is a appropriate sequence of positive real numbers. We point out that the advantage of this procedure is that it is very easy to consider the same empirical procedure to select the optimal $V_{\mathbf{k}}$.
How to select the bandwidths parameters in practice? It is well known that the question of the smoothing parameters is very crucial for the performance of the estimators in the kernel method. Usually, the ideal theoretical choices is obtained by minimizing the explicit quadratic error. The asymptotic normality given in Theorem 2 is a basic ingredient to determine the

leading term of this error. Indeed, the result of Theorem 2 implies that, under suitable condition we have the asymptotic bias and variance

$$B_{\mathbf{n}}^2(x) + \frac{\sigma^2(x)}{\widehat{\mathbf{n}} h_H^3 \phi_x(h_K)}.$$

Then, the optimal smoothing parameters is solution of

$$\min_{h_H, h_K} B_H^2 h_H^2 + B_K^2 h_K^2 + \frac{\sigma^2(x)}{\widehat{\mathbf{n}} h_H^3 \phi_x(h_K)}.$$

However, the practical utilization of this criterium requires some additional computational efforts. More precisely, it requires the estimation of the unknown quantities, Φ_0' and $f^x(\theta(x))$. Clearly, all these estimations can be obtained by using a pilots estimators of the the conditional density $f^x(y)$. Such estimations is possible by using the kernel methods, with an adaptation to the spatial case, the bandwidths selectors studied by Laksaci *et al.* (2011) as follows

$$(8) \qquad CVPDF = \frac{1}{\widehat{\mathbf{n}}} \sum_{\mathbf{i} \in I_{\mathbf{n}}} W_1(X_{\mathbf{i}}) \int \widehat{f}^{X_{\mathbf{i}}^{-\mathbf{i}}}{}^2(y) W_2(y) dy - \frac{2}{\widehat{\mathbf{n}}} \sum_{\mathbf{i} \in I_{\mathbf{n}}} \widehat{f}^{X_{\mathbf{i}}^{-\mathbf{i}}}(Y_{\mathbf{i}}) W_1(X_{\mathbf{i}}) W_2(Y_{\mathbf{i}})$$

where

$$\widehat{f}^{X_{\mathbf{i}}^{-\mathbf{i}}}(y) = \frac{h_H^{-1} \sum_{\mathbf{j} \in I_{\mathbf{n},\varsigma_{\mathbf{n}}}^{\mathbf{i}}} K(h_K^{-1} d(X_{\mathbf{i}}, X_{\mathbf{j}})) H'(h_H^{-1}(y - Y_{\mathbf{j}}))}{\sum_{\mathbf{j} \in I_{\mathbf{n},\varsigma_{\mathbf{n}}}^{\mathbf{i}}} K(h_K^{-1} d(X_{\mathbf{i}}, X_{\mathbf{j}}))}$$

with

$$I_{\mathbf{n},\varsigma_{\mathbf{n}}}^{\mathbf{i}} = \{\mathbf{j} \text{ such that } \|\mathbf{j} - \mathbf{i}\| \geq \varsigma_{\mathbf{n}} \}.$$

Nevertheless, a data-driven method allows to overcome this additional computation is very important in practice. Although, this problem is not the subject of this paper, a reasonable method is to consider the spatial version of the method used in Ferraty and Vieu (2006).

$$(9) \qquad (h_K^{opt}, h_H^{opt}) = \arg\min_{h_H, h_K} \left\{ \frac{1}{\widehat{\mathbf{n}}} \sum_{\mathbf{i} \in I_{\mathbf{n}}} (Y_{\mathbf{i}} - \widehat{\theta}^{-\mathbf{i}}(X_{\mathbf{i}}))^2 \right\}$$

where

$$\widehat{\theta}^{-\mathbf{i}}(X_{\mathbf{i}}) = \arg\max_{y \in \mathbb{R}} \widehat{f}^{X_{\mathbf{i}}^{-\mathbf{i}}}(y).$$

The asymptotic optimality of this method is one of the natural prospects of the present work. *Confidence intervals :* The main application of Theorem 2 is to build confidence band for the true value of $\theta(x)$. To do that, we must estimate the quantity $\sigma_1^2(x)$. The latter can be estimated by estimating empirically β_1 and β_2 by

$$\widehat{\beta}_1 = \frac{1}{\widehat{\mathbf{n}} \phi_x(h_K)} \sum_{\mathbf{i} \in I_{\mathbf{n}}} K(h_K^{-1} d(x, X_{\mathbf{i}})) \quad \text{and} \quad \widehat{\beta}_2 = \frac{1}{\widehat{\mathbf{n}} \phi_x(h_K)} \sum_{\mathbf{i} \in I_{\mathbf{n}}} K^2(h_K^{-1} d(x, X_{\mathbf{i}})).$$

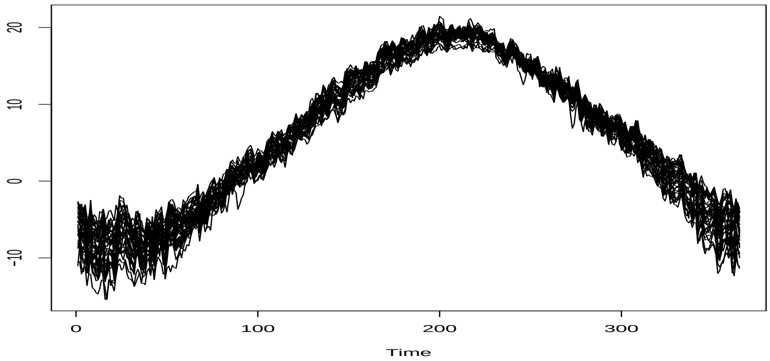


FIGURE 3.1 – daily temperature in 35 weather stations

Clearly, with this, estimation the function $\phi_x(\cdot)$ does not appear in the calculation of the confidence interval by simplification. More precisely, we obtain the following approximate $(1-\zeta)$ confidence band for $\theta(x)$

$$\widehat{\theta}(x) \pm t_{1-\zeta/2} \times \left(\frac{\widehat{\sigma_1^2(x)}}{\widehat{\mathbf{n}} h_H \phi_x(h_K)} \right)^{1/2}$$

where $t_{1-\zeta/2}$ denotes the $1-\zeta/2$ quantile of the standard normal distribution.

3.4.2 Real data application

In this section, we are interested in how our procedure works in practice with special attention on the influence of the choice of the different parameters on the efficiency of the estimate. Our main purpose in this illustrative application is to predict, via the conditional mode estimation, the temperature of one month given the daily temperature of the others months. The data set used here is the *maritimes-data* of *geofd package* in R (see Fig. 3.1) which contains the daily temperature averaged over 1960-1994 observed at 35 Canadian Maritime weather stations. It should be noted that, the spatial prediction of meteorological data is of interest, in particular, for studying the microclimate conditions in mountainous terrain, resource management or calibration of satellite sensors. Recently, this problem has been modeled by cokriging method (see, Deldico 2009). It is well known that in practice it is very hard to have some information on the shape of the relationship between the functional variable and the scalar response because the so-called *scatterplot* which is a graphical tool for exploring the relationship between the explanatory variables and the scalar response is not available in functional statistics. So, we can say that the nonparametric modelization is more appropriate for this kind of data. In this context, the conditional mode estimation provides alternative approach to regression method and gives better result than the classical

mean regression in several situations. So, it Is very important to test the efficiency of this model as spatial prediction in meteorological data.
Now, in order to exploit the spatial correlation of the data set and to determine its influence in prediction analysis we examine the performance of our methodology over four separates months (January, March, June and September) by using the two estimation procedures discussed in the previous section. Specifically, for each day t in month $m = 1, 3, 6, 9$ we predict the daily temperature in some fixed station $\mathbf{k}$ by :

$$\widehat{\theta}_{Method1}(X_{\mathbf{k}}^{t,m}) = \arg\min \frac{h_H^{-1} \sum_{\mathbf{i}=1}^{34} K(h_K^{-1} d(X_{\mathbf{k}}^{t,m}, X_{\mathbf{i}}^{t,m})) H(h_H^{-1}(y - Y_{\mathbf{i}}^{t,m})) \mathbb{1}_{W_{\mathbf{k}}}(\mathbf{i})}{\sum_{\mathbf{i}=1}^{34} K(h_K^{-1} d(X_{\mathbf{k}}^{t,m}, X_{\mathbf{i}}^{t,m})) \mathbb{1}_{W_{\mathbf{k}}}(\mathbf{i})}$$

and

$$\widehat{\theta}_{Method2}(X_{\mathbf{k}}^{t,m}) = \arg\min \frac{h_H^{-1} \sum_{\mathbf{i}=1}^{34} K(h_K^{-1} d(X_{\mathbf{k}}^{t,m}, X_{\mathbf{i}}^{t,m})) H(h_H^{-1}(y - Y_{\mathbf{i}}^{t,m})) \mathbb{1}_{V_{\mathbf{k}}}(\mathbf{i})}{\sum_{\mathbf{i}=1}^{34} K(h_K^{-1} d(X_{\mathbf{k}}^{t,m}, X_{\mathbf{i}}^{t,m})) \mathbb{1}_{V_{\mathbf{k}}}(\mathbf{i})}$$

where $W_{\mathbf{k}}$ and $V_{\mathbf{k}}$ are defined respectively by (6) and (7) and $X_{\mathbf{i}}^{t,m}$ is the curve of the daily temperature at station $\mathbf{i}$ by leaving the temperature of day t in month m which is considered as response variable $Y_{\mathbf{i}}^{t,m}$.
For this comparison study, we treat the both estimators with the same conditions. Indeed, the smoothing parameters (h_K, h_H) are locally chosen by cross-validation on the k-nearest neighbors with respect to the criterium (9). We point out that the quantities $\varsigma_{\mathbf{n}}$ and $\nu_{\mathbf{n}}$ are optimally selected over the nearest neighbors locations with respect to the Euclidean norm on the geographic coordinates while $\iota_{\mathbf{n}}$ is selected among the quantile of order q of the estimated vector trace-variogramme $\gamma(\mathbf{k}, \mathbf{i})$ given by the code *okfd in R*. Concerning the semi-metric choice we consider the semi-metric based on the m first eigenfunctions of the empirical covariance operator associated to the m greatest eigenvalues (see Ferraty and Vieu 2006, pp. 28 and 223 for more discussion). This choice is motivated by the non-differentiability of the curves $X_{\mathbf{i}}$. The semi-metric parameter m is chosen from the training sample by the cross-validation method. For all these parameters we use the mean squared prediction errors (MSE), defined by the following quantities :

$$MSE(m) = \frac{1}{31} \sum_{t=1}^{31} \left(Y_{\mathbf{k}}^{t,m} - \widehat{\theta}_{Method.}(X_{\mathbf{k}}^{t,m}) \right)^2$$

as accuracy criterion. Finally, we precise that we use quadratic kernels on $(-1, 1)$ $(K = H)$. The results are given in the following graphs where we draw the curves corresponding to the observed values (dashed curve) and estimated values (solid curve) for the different months. In Figure 3.2 we draw the results of the first method and the results of the second method is given in Figure 3.3
We observe that the first method gives the best results regarding the largest MSE equals to 0.636 while the smallest MSE in the second method is 0.74. But generally the both methods give a satisfactory level of accuracy.

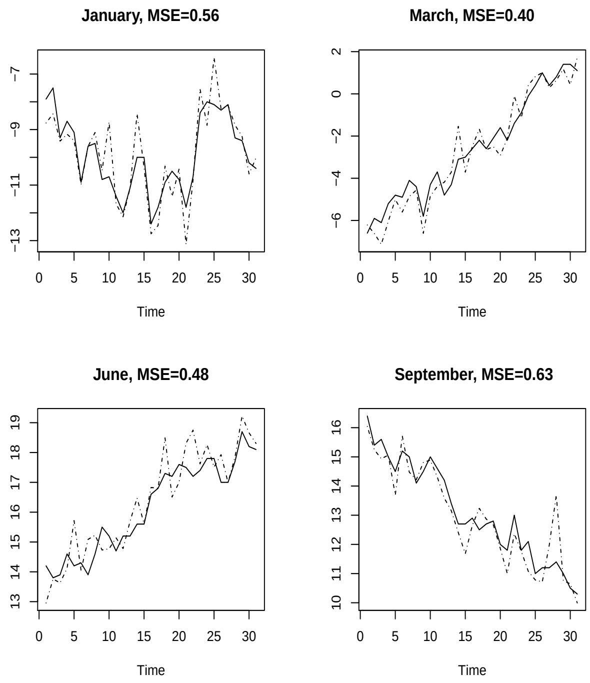


FIGURE 3.2 – The prediction results with vicinity set defined by trace-variogramme function

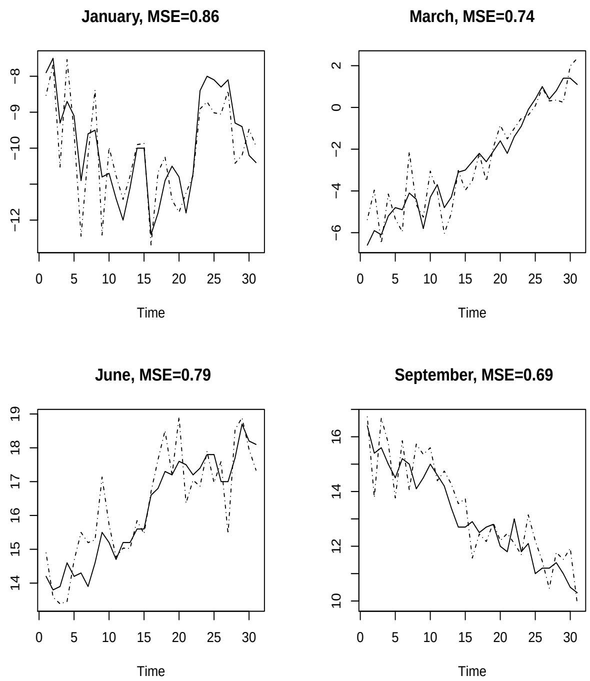


FIGURE 3.3 – The prediction results with vicinity set defined by locations distance

3.5 Appendix

We first state the following lemmas which are due to Carbon et al. (1997). They are needed for the convergence of our estimates. There proofs will then be omitted.

Lemma 3.5.1 *Suppose $E_1, ..., E_r$ be sets containing m sites each with $dist(E_i,\ E_j) \geq \gamma$ for all $i \neq j$ where $1 \leq i \leq r$ and $1 \leq j \leq r$. Suppose $Z_1, ..., Z_r$ is a sequence of real-valued r.v.'s measurable with respect to $\mathcal{B}(E_1), ..., \mathcal{B}(E_r)$ respectively, and Z_i takes values in $[a, b]$. Then there exists a sequence of independent r.v.'s $Z_1^*, ..., Z_r^*$ independent of $Z_1, ..., Z_r$ such that Z_i^* has the same distribution as Z_i and satisfies*

$$\sum_{i=1}^{r} E|Z_i - Z_i^*| \leq 2r(b-a)\psi((r-1)m, m)\varphi(\gamma).$$

Lemma 3.5.2
(i) Suppose that (3) holds. Denote by $\mathcal{L}_r(\mathcal{F})$ the class of $\mathcal{F}-$measurable r.v.'s X satisfying $\|X\|_r = (E|X|^r)^{1/r} < \infty$. Suppose $X \in \mathcal{L}_r(\mathcal{B}(E))$ and $Y \in \mathcal{L}_s(\mathcal{B}(E'))$. Assume also that $1 \leq r,\ s,\ t < \infty$ and $r^{-1} + s^{-1} + t^{-1} = 1$. Then

$$(10) \qquad |EXY - EXEY| \leq C\|X\|_r\|Y\|_s\{\psi(Card(E), Card(E'))\varphi(dist(E, E'))\}^{1/t}.$$

(ii) For r.v.'s bounded with probability 1, the right-hand side of (9) can be replaced by

$$C\psi(Card(E), Card(E'))\varphi(dist(E, E')).$$

Proof of Lemma 3.2.1 : We have

$$\left\|\widehat{f}_N^{x\,(1)}(\theta(x)) - E\left[\widehat{f}_N^{x\,(1)}(\theta(x))\right]\right\|_p = \frac{1}{\widehat{\mathbf{n}}\, h_H^2 E K_{\mathbf{1}}}\left\|\sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} \theta_\mathbf{i}\right\|_p$$

where

$$\theta_\mathbf{i} = K_\mathbf{i} H_\mathbf{i}^{(1)}(\theta(x))) - E\left[K_\mathbf{i} H_\mathbf{i}^{(1)}(\theta(x)))\right].$$

We have $EK_\mathbf{1} = O(\phi_x(h_K))$, (because of H3), so it remains to show that

$$\left\|\sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} \theta_\mathbf{i}\right\|_p = O(\sqrt{\widehat{\mathbf{n}} h_H \phi_x(h_K)}).$$

Clearly, if we show the claimed result for $p = 2r$ we get the others lower values of p by the Holder inequality. Thus, it suffices to proof the case $p = 2r$. To do that we use the same ideas of Gao *et al.* (2008) and Abdi et al (2010). Indeed, we have

$$E\left[\left(\sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} \theta_\mathbf{i}\right)^{2r}\right] = \sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} E\left[\theta_\mathbf{i}^{2r}\right] + \sum_{s=1}^{2r-1} \sum_{\nu_0+\nu_1+...+\nu_s=2r} V_s(\nu_0, \nu_1, ..., \nu_s)$$

where $\sum_{\nu_0+\nu_1+...+\nu_s=2r}$ is the summation over $(\nu_0,\nu_1,...,\nu_s)$ with positive integer components satisfying $\nu_0+\nu_1+...+\nu_s=2r$ and

$$V_s(\nu_0,\nu_1,...,\nu_s)=\sum_{\mathbf{i}_0\neq\mathbf{i}_1\neq...\neq\mathbf{i}_s\in\mathcal{I}_\mathbf{n}} E\left[\theta_{\mathbf{i}_0}^{\nu_0}\theta_{\mathbf{i}_1}^{\nu_1}...\theta_{\mathbf{i}_s}^{\nu_s}\right].$$

Firstly, by stationarity, we have

$$\sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} E\left(\theta_\mathbf{i}\right)^{2r}\leq C\widehat{\mathbf{n}}E\left(|\theta_\mathbf{i}|\right)^{2r}\leq \widehat{\mathbf{n}}E\left(K_\mathbf{i}H_\mathbf{i}^{(1)}(\theta(x))\right)^{2r}\leq C\widehat{\mathbf{n}}h_H\phi_x(h_K).$$

Now, we control the second term $V_s(\nu_0,\nu_1,...,\nu_s)$. For this, we distingue the two cases $s<r$ and $s\geq r$. The second case can be evaluated by straightforward modification of the proof of Lemma 3.4 of Gao *et al.* (2008) by taking into account, here,

$$E\left|\theta_{\mathbf{i}_0}^{\nu_0}\theta_{\mathbf{i}_2}^{\nu_2}...\theta_{\mathbf{i}_s}^{\nu_s}\right|\leq Ch_H^{1+s}\phi_x^{1+v_s}(h_K).$$

It follows that, for a certain a real sequence $P:=P_\mathbf{n}$ we have

$$V_s(\nu_0,\nu_1,...,\nu_s)\leq C(\widehat{\mathbf{n}})^r\left(P^{Nr}h_H^{1+s}\phi_x(h_K)^{(1+v_s)}+P^{Nr-1-\delta}\right).$$

So, if we take $P=\phi_x(h_K)^{-(1+v_{Nr})/1+\delta)}$, we obtain that

$$V_s(\nu_0,\nu_1,...,\nu_s)=O\left((\widehat{\mathbf{n}}h_H\phi_x(h_K))^r\right),\qquad\text{for}\qquad r\leq s\leq 2r-1.$$

Next, the case $s>r$ is evaluated by the same way as in Lemma 3.3 in Gao *et al.* (2008). Indeed, we denote by

$$V_{s1}=\sum_{\mathbf{i}_0\neq\mathbf{i}_1\neq...\neq\mathbf{i}_s}\left[E\left(\prod_{j=0}^{s}\theta_{\mathbf{i}_j}^{\nu_j}\right)-\prod_{j=0}^{s}E\theta_{\mathbf{i}_j}^{\nu_j}\right]\text{ and }V_{s2}=\sum_{\mathbf{i}_0\neq\mathbf{i}_1\neq...\neq\mathbf{i}_s}\prod_{j=0}^{s}E\theta_{\mathbf{i}_j}^{\nu_j}.$$

So,

$$V_s(\nu_0,\nu_1,...,\nu_s)=V_{s1}+V_{s2}.$$

It is clear that,

$$|V_{s2}|\leq C\left(\widehat{\mathbf{n}}h_H\phi_x(h_K)\right)^{s+1}.$$

Furthermore for the term V_{s1}, we have

$$E\left(\prod_{j=0}^{s}\theta_{\mathbf{i}_j}^{\nu_j}\right)-\prod_{j=0}^{s}E\theta_{\mathbf{i}_j}^{\nu_j}=\sum_{l=0}^{s-1}\left(\prod_{j=0}^{l-1}E\theta_{\mathbf{i}_j}^{\nu_j}\right)\left(E\left[\prod_{j=l}^{s}\theta_{\mathbf{i}_j}^{\nu_j}\right]-E\left[\theta_{\mathbf{i}_l}^{\nu_l}\right]E\left[\prod_{j=l+1}^{s}\theta_{\mathbf{i}_j}^{\nu_j}\right]\right)$$

where we define $\prod_{j=l}^{s}\cdot=1$ if $l>s$. Then we obtain

$$|V_{s1}|\leq\sum_{l=0}^{s-1}\left(\widehat{\mathbf{n}}h_H\phi_x(h_K)\right)^l\sum_{\mathbf{i}_l\neq...\neq\mathbf{i}_s}\left|E\left[\prod_{j=l}^{s}\xi_{\mathbf{i}_j}^{\nu_j}\right]-E\left[\xi_{\mathbf{i}_l}^{\nu_l}\right]E\left[\prod_{j=l+1}^{s}\xi_{\mathbf{i}_j}^{\nu_j}\right]\right|=\sum_{l=0}^{s-1}\left(\widehat{\mathbf{n}}h_H\phi_x(h_K)\right)^lV_{ls1}.$$

Let P be some positive real, we have

$$\begin{aligned} V_{ls1} &= \sum_{0<dist(\{\mathbf{i}_l\},\{\mathbf{i}_{l+1},\ldots,\mathbf{i}_s\})\leq P} \left| E\left[\prod_{j=l}^{s} \theta_{\mathbf{i}_j}^{\nu_j}\right] - E\left[\theta_{\mathbf{i}_l}^{\nu_l}\right] E\left[\prod_{j=l+1}^{s} \theta_{\mathbf{i}_j}^{\nu_j}\right]\right| \\ &\quad + \sum_{0<dist(\{\mathbf{i}_l\},\{\mathbf{i}_{l+1},\ldots,\mathbf{i}_s\})> P} \left| E\left[\prod_{j=l}^{s} \theta_{\mathbf{i}_j}^{\nu_j}\right] - E\left[\theta_{\mathbf{i}_l}^{\nu_l}\right] E\left[\prod_{j=l+1}^{s} \theta_{\mathbf{i}_j}^{\nu_j}\right]\right| \\ &\quad := V_{ls11} + V_{ls12}. \end{aligned}$$

Once again by (2) and (H2) we have

$$\left| E\left[\prod_{j=l}^{s} \theta_{\mathbf{i}_j}^{\nu_j}\right] - E\left[\theta_{\mathbf{i}_l}^{\nu_l}\right] E\left[\prod_{j=l+1}^{s} \theta_{\mathbf{i}_j}^{\nu_j}\right]\right| \leq C h_H^{s-l} \phi_x(h_K)^{1+v_{s+1-l}}.$$

Thus, we have

$$V_{ls11} \leq C h_H^{s-l} \phi_x(h_K)^{1+v_{s+1-l}} \widehat{\mathbf{n}}^{(s-l)} P^N.$$

Since the variables $\theta_{\mathbf{i}}$ are bounded, we have (see Lemma 4.5.2)

$$V_{ls12} \leq C \widehat{\mathbf{n}}^{(s-l)} \sum_{t=P+1}^{\infty} t^{N-1} \varphi(t).$$

Combining the upper bounds of V_{ls11} and V_{ls12}, we have

$$\begin{aligned} |V_{s1}| &\leq C \sum_{l=0}^{s-1} \left(\widehat{\mathbf{n}} h_H \phi_x(h_K)\right)^l \left[h_H^{s-l} \phi_x(h_K)^{1+v_{s+1-l}} \widehat{\mathbf{n}}^{(s-l)} P^N + \widehat{\mathbf{n}}^{(s-l)} \sum_{t=P+1}^{\infty} t^{N-1} \varphi(t) \right] \\ &\leq C \left(\widehat{\mathbf{n}} h_H \phi_x(h_K)\right)^{(s+1)} \\ &\qquad \sum_{l=0}^{s-1} \left(\widehat{\mathbf{n}} h_H \phi_x(h_K)\right)^{l-s-1} \left[h_H^{s-l} \phi_x(h_K)^{1+v_{s+1-l}} \widehat{\mathbf{n}}^{(s-l)} P^N + \widehat{\mathbf{n}}^{(s-l)} \sum_{t=P+1}^{\infty} t^{N-1} \varphi(t) \right]. \end{aligned}$$

Taking $P = (h_H \phi_x(h_K))^{-1/N}$, we obtain

$$|V_{s1}| \leq C \left(\widehat{\mathbf{n}} h_H \phi_x(h_K)\right)^{(s+1)}$$

.

Proof of Lemma 3.2.2
It is easy to see that

$$E\left[\widehat{f}_N^{x\,(1)}(\theta(x))\right] - f^{x\,(1)}(\theta(x)) = \frac{1}{h_H^2 E K_1} E\left[K_1 E\left[H_1^{(1)}(\theta(x)) \mid X_1\right]\right] - f^{x\,(1)}(\theta(x)).$$

By an integration by part and the change of variables $t = \dfrac{y-z}{h_H}$, we have

$$E\left[\widehat{f}_N^{x\,(1)}(\theta(x))\right] - f^{x\,(1)}(\theta(x)) \leq \frac{1}{EK_1}\left(EK_1 \int H(t)\left|f^{X_1\,(1)}(\theta(x)-h_H t) - f^{x\,(1)}(\theta(x))\right| dt\right).$$

Hypotheses (H4) and (H6) allow to get the desired result.

+

Proof of Lemma 3.2.3. It is clear that, for all $\epsilon < 1$, we have

$$P\left(\widehat{f}_D^x = 0\right) \leq P\left(\widehat{f}_D^x \leq 1-\epsilon\right) \leq P\left(|\widehat{f}_D^x - E[\widehat{f}_D^x]| \geq \epsilon\right).$$

The Markov's inequality allows to get, for any $p > 0$,

$$P\left(|\widehat{f}_D^x - E[\widehat{f}_D^x]| \geq \epsilon\right) \leq \frac{E\left[|\widehat{f}_D^x - E[\widehat{f}_D^x]|^p\right]}{\epsilon^p}.$$

So

$$\left(P\left(\widehat{F}_D^x = 0\right)\right)^{1/p} = O\left(\left\|\widehat{f}_D^x - E[\widehat{f}_D^x]\right\|_p\right).$$

The computation of $\left\|\widehat{f}_D^x - E[\widehat{f}_D^x]\right\|_p$ can be done by following the same arguments as those invoked to get (5). This yields the proof.

Proof of Lemma 4.4.6 We start by writing $\forall y \in S$

$$E[\widehat{f}_N^{x\,(1)}(y)] = \frac{1}{E[K_{\mathbf{1}}]} E\left[K_{\mathbf{1}} E[h_H^{-2} H_{\mathbf{1}}'|X]\right] \text{ with } h_H^{-2} E\left[H_{\mathbf{1}}'|X\right] = \int_{\mathbb{R}} H(t) f^{X\,(1)}(y-h_H t)dt.$$

Next, using a Taylor expansion under (H4), as follows

$$h_H^{-2} E\left[H_{\mathbf{1}}'|X\right] = f^{X\,(1)}(y) + \frac{h_H}{2}\left(\int tH(t)dt\right)\frac{\partial^2 f^X(y)}{\partial^2 y} + o(h_H).$$

Thus, we get

$$E\left[\widehat{f}_N^{x\,(1)}(y)\right] = \frac{1}{E[K_{\mathbf{1}}]}\left(E\left[K_{\mathbf{1}} f^{X\,(1)}(y)\right] + h_H\left(\int tH(t)dt\right) E\left[K_{\mathbf{1}}\frac{\partial^2 f^X(y)}{\partial^2 y}\right] + o(h_H)\right).$$

Noting $\psi_l(\cdot, y) := \frac{\partial^l f^{\cdot}(y)}{\partial^l y}$: for $l \in \{1, 2\}$, since $\Phi_l(0) = 0$, we have

$$\begin{aligned} E\left[K_{\mathbf{1}}\psi_l(X,y)\right] &= \psi_l(x,y)E[K_{\mathbf{1}}] + E\left[K_{\mathbf{1}}\left(\psi_l(X,y) - \psi_l(x,y)\right)\right] \\ &= \psi_l(x,y)E[K_{\mathbf{1}}] + E\left[K_{\mathbf{1}}\left(\Phi_l(d(x,X))\right)\right] \\ &= \psi_l(x,y)E[K_{\mathbf{1}}] + \Phi_l'(0)E\left[d(x,X)K_{\mathbf{1}}\right] + o(E\left[d(x,X)K_{\mathbf{1}}\right]). \end{aligned}$$

Therefore,

$$\begin{aligned} E\left[\widehat{f}_N^{x\,(1)}(y)\right] &= f^{x\,(1)}(y) + \frac{h_H}{2}\frac{\partial^2 f^x(y)}{\partial y^2}\int tH(t)dt + o\left(h_H^2\frac{E\left[d(x,X)K_{\mathbf{1}}\right]}{E[K_{\mathbf{1}}]}\right) \\ &\quad + \Phi_0'(0)\frac{E\left[d(x,X)K_{\mathbf{1}}\right]}{E[K_{\mathbf{1}}]} + o\left(\frac{E\left[d(x,X)K_{\mathbf{1}}\right]}{E[K_{\mathbf{1}}]}\right). \end{aligned}$$

By using the same idea's of Ferraty *et al.* (2007) we obtain that

$$\frac{1}{\phi_x(h_K)}E\left[d(x,X)K_{\mathbf{1}}\right] = h_K\left(K(1) - \int_0^1 (sK(s))'\beta_x(s)ds + o(1)\right)$$

and

$$\frac{1}{\phi_x(h_K)}E\left[K_{\mathbf{1}}\right] = K(1) - \int_0^1 K'(s)\beta_x(s)ds + o(1).$$

So,

$$\begin{aligned} E\left[\widehat{f}_N^{x\,(1)}(y)\right] &= f^{x\,(1)}(y) + \frac{h_H}{2}\frac{\partial^2 f^x(y)}{\partial y^2}\int tH(t)dt \\ &\quad + h_K\Phi_0'(0)\frac{\left(K(1) - \int_0^1 (sK(s))'\beta_x(s)ds\right)}{\left(K(1) - \int_0^1 K'(s)\beta_x(s)ds\right)} + o(h_H) + o(h_K). \end{aligned}$$

In particular for $y = \theta(x)$ we obtain the desired result of Lemma.
Proof of Lemma 4.4.7 For the variance term $Var[\widehat{f}_N^x(y)]$, we have

$$Var[\widehat{f}_N^{x\,(1)}(y)] = \frac{1}{(h_H^2\widehat{\mathbf{n}}E\left[K_{\mathbf{1}}\right])^2}Var\left[\sum_{\mathbf{i}\in I_{\mathbf{n}}} D_{\mathbf{i}}\right]$$

where

$$D_{\mathbf{i}} = K_{\mathbf{i}}H_{\mathbf{i}}' - E\left[K_{\mathbf{i}}H_{\mathbf{i}}'\right].$$

Thus

$$Var[\widehat{f}_N^{x\,(1)}(y)] = \frac{1}{\widehat{\mathbf{n}}(h_H^2 E\left[K_{\mathbf{1}}\right])^2}Var\left[D_{\mathbf{1}}\right] + \frac{1}{(h_H^2\widehat{\mathbf{n}}E\left[K_{\mathbf{1}}\right])^2}\sum_{\mathbf{i}\neq\mathbf{j}} Cov(D_{\mathbf{i}}, D_{\mathbf{j}}).$$

Let us calculate the quantity $Var\left[D_{\mathbf{1}}\right]$. We have :

$$\begin{aligned} Var\left[D_{\mathbf{1}}\right] &= E\left[K_{\mathbf{1}}^2 H_{\mathbf{1}}'^2\right] - \left(E\left[K_{\mathbf{1}}H_{\mathbf{1}}'\right]\right)^2 \\ &= E\left[K_{\mathbf{1}}^2\right]\frac{E\left[K_{\mathbf{1}}^2 H_{\mathbf{1}}'^2\right]}{E\left[K_{\mathbf{1}}^2\right]} - \left(E\left[K_{\mathbf{1}}\right]\right)^2\left(\frac{E\left[K_{\mathbf{1}}H_{\mathbf{1}}'\right]}{E\left[K_{\mathbf{1}}\right]}\right)^2. \end{aligned}$$

So, by using the same arguments as those used in the previous lemma we get

$$\begin{aligned}\frac{1}{\phi_x(h_K)}E\left[K_{\mathbf{1}}^2\right] &= K^2(1)-\int_0^1 (K^2(s))'\beta_x(s))ds+o(1)\\ \frac{E\left[K_{\mathbf{1}}^2H_{\mathbf{1}}'^2\right]}{E\left[K_{\mathbf{1}}^2\right]} &= h_H f^x(y)\int H'^2(t)dt+o(h_H)\\ \frac{E[K_{\mathbf{1}}H_{\mathbf{1}}']}{E\left[K_{\mathbf{1}}\right]} &= h_H f^x(y)\int H'(t)dt+o(h_H)\end{aligned}$$

which implies that
(11)

$$Var\left[D_{\mathbf{i}}\right]=h_H\phi_x(h_K)f^x(y)\int H'^2(t)dt\left(K^2(1)-\int_0^1 (K^2(s))'\beta_x(s))ds\right)+o\left(h_H\phi_x(h_K)\right).$$

Now, let us focus on the covariance term. To do that, we define

$$E_1=\{\mathbf{i},\mathbf{j}\in\mathcal{I}_{\mathbf{n}}:0<\|\mathbf{i}-\mathbf{j}\|\le c_{\mathbf{n}}\}\ \text{ and }\ E_2=\{\mathbf{i},\mathbf{j}\in\mathcal{I}_{\mathbf{n}}:\|\mathbf{i}-\mathbf{j}\|>c_{\mathbf{n}}\}.$$

For all $(\mathbf{i},\mathbf{j})\in E_1^2$ we write

$$Cov\left(D_{\mathbf{i}},D_{\mathbf{j}}\right) \;=\; E\left[K_{\mathbf{i}}K_{\mathbf{j}}H_{\mathbf{i}}'H_{\mathbf{j}}'\right]-\left(E\left[K_{\mathbf{i}}H_{\mathbf{i}}'\right]\right)^2$$

and we use the fact that $\mathrm{E}\left[H_{\mathbf{i}}'H_{\mathbf{j}}'|(X_{\mathbf{i}},X_{\mathbf{j}})\right]=O(h_H^2);\ \forall\,\mathbf{i}\neq\mathbf{j},\mathrm{E}\left[H_{\mathbf{i}}'|X_{\mathbf{i}}\right]=O(h_H);\ \forall\,\mathbf{i}$, under (H2) and (H5) we get

$$E\left[K_{\mathbf{i}}K_{\mathbf{j}}H_{\mathbf{i}}'H_{\mathbf{j}}'\right]\le Ch_H^2P\left[(X_{\mathbf{i}},X_{\mathbf{j}})\in B(x,h_K)\times B(x,h_K)\right]$$

and

$$E\left[K_{\mathbf{i}}H_{\mathbf{i}}'\right]\le Ch_HP\left(X_{\mathbf{i}}\in B(x,h_K)\right).$$

It follows that

$$Cov\left(D_{\mathbf{i}},D_{\mathbf{j}}\right)\le Ch_H^2\phi_x(h_K)(\phi_x(h_K)+\phi_x^{v_1}(h_K)).$$

So

$$\sum_{E_1}Cov\left(D_{\mathbf{i}},D_{\mathbf{j}}\right)\le C\widehat{\mathbf{n}}c_{\mathbf{n}}^Nh_H^2\phi_x^{1+v_1}(h_K).$$

On the other hand, Lemma 4.5.2 and $|D_{\mathbf{i}}|\le C$, permit to write that $(\mathbf{i},\mathbf{j})\in E_2^2$

$$\left|Cov\left(D_{\mathbf{i}},D_{\mathbf{j}}\right)\right|\le C\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)$$

and

$$\begin{aligned}\sum_{E_2}Cov\left(D_{\mathbf{i}},D_{\mathbf{j}}\right) &\le C\sum_{E_2}\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)\\ &\le C\widehat{\mathbf{n}}\sum_{\mathbf{i}:\|\mathbf{i}\|>c_{\mathbf{n}}}\varphi\left(\|\mathbf{i}\|\right)\\ &\le C\widehat{\mathbf{n}}c_{\mathbf{n}}^{-Na}\sum_{\mathbf{i}:\|\mathbf{i}\|>c_{\mathbf{n}}}\|\mathbf{i}\|^{Na}\varphi\left(\|\mathbf{i}\|\right).\end{aligned}$$

Finally, we have :

$$\sum_{\mathbf{i}\neq\mathbf{j}} Cov(D_{\mathbf{i}}, D_{\mathbf{j}}) \leq \left(C\widehat{\mathbf{n}} c_{\mathbf{n}}^N h_H^2 \phi_x^{1+v_1}(h_K) + C\widehat{\mathbf{n}} c_{\mathbf{n}}^{-Na} \sum_{\mathbf{i}:\|\mathbf{i}\|>c_{\mathbf{n}}} \|\mathbf{i}\|^{Na} \varphi\left(\|\mathbf{i}\|\right) \right).$$

Let $c_{\mathbf{n}} = (h_H^{2/(a+1)} \phi_x^{1/a}(h_K))^{-1/N}$, then we obtain that

$$\sum Cov\left(D_{\mathbf{i}}, D_{\mathbf{j}}\right) = o\left(\widehat{\mathbf{n}} h_H \phi_x(h_K)\right).$$

In conclusion, we have

$$Var[\widehat{f}_N^{x\,(1)}(y)] = \frac{f^x(y)}{\widehat{\mathbf{n}} h_H^3 \phi_x(h_K)} \left(\int H'^2(t)dt \right) \left(\frac{\left(K^2(1) - \int_0^1 (K^2(s))'\beta_x(s)ds\right)}{\left(K(1) - \int_0^1 K'(s)\beta_x(s)ds\right)^2} \right) + o\left(\frac{1}{\widehat{\mathbf{n}} h_H^3 \phi_x(h_K)} \right)$$

The result of the lemma corresponds to the case $y = \theta(x)$.

Proof of Lemma 3.3.3

We set

$$S_{\mathbf{n}} = \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \Lambda_{\mathbf{i}}$$

where

$$\Lambda_{\mathbf{i}} := \frac{\sqrt{h_H \phi_x(h_K)}}{h_H E[K_{\mathbf{1}}]} D_{\mathbf{i}}. \tag{12}$$

Clearly , we have

$$\sqrt{\widehat{\mathbf{n}} h_H^3 \phi_x(h_K)}\, [\sigma(x)]^{-1} \left(\widehat{f}_N^{x\,(1)}(y) - E\widehat{f}_N^{x\,(1)}(y) \right) = \left(\widehat{\mathbf{n}}(\sigma^2(x)) \right)^{-1/2} S_{\mathbf{n}}.$$

So, the asymptotic normality of $(\widehat{\mathbf{n}}(\sigma(x))^2)^{-1/2} S_{\mathbf{n}}$ is sufficient to show the proof of this Lemma. The latter last is shown by the blocking method, where the random variables $\Lambda_{\mathbf{j}}$ are grouped into blocks of different sizes defined as follows

$$W(1, \mathbf{n}, x, \mathbf{j}) = \sum_{i_k = j_k(p_{\mathbf{n}}+q_{\mathbf{n}})+1,\ k=1,\ldots,N}^{j_k(p_{\mathbf{n}}+q_{\mathbf{n}})+p_{\mathbf{n}}} \Lambda_{\mathbf{i}},$$

$$W(2, \mathbf{n}, x, \mathbf{j}) = \sum_{i_k = j_k(p_{\mathbf{n}}+q_{\mathbf{n}})+1,\ k=1,\ldots,N-1}^{j_k(p_{\mathbf{n}}+q_{\mathbf{n}})+p_{\mathbf{n}}} \quad \sum_{i_N = j_N(p_{\mathbf{n}}+q_{\mathbf{n}})+p_{\mathbf{n}}+1}^{(j_N+1)(p_{\mathbf{n}}+q_{\mathbf{n}})} \Lambda_{\mathbf{i}},$$

$$W(3,\mathbf{n},x,\mathbf{j})=\sum_{i_k=j_k(p_\mathbf{n}+q_\mathbf{n})+1,\ k=1,\dots,N-2}^{j_k(p_\mathbf{n}+q_\mathbf{n})+p_\mathbf{n}}\ \sum_{i_{N-1}=j_{N-1}(p_\mathbf{n}+q_\mathbf{n})+p_\mathbf{n}+1}^{(j_{N-1}+1)(p_\mathbf{n}+q_\mathbf{n})}\ \sum_{i_N=j_N(p_\mathbf{n}+q_\mathbf{n})+1}^{j_N(p_\mathbf{n}+q_\mathbf{n})+p_\mathbf{n}}\Lambda_\mathbf{i},$$

$$W(4,\mathbf{n},x,\mathbf{j})=\sum_{i_k=j_k(p_\mathbf{n}+q_\mathbf{n})+1,\ k=1,\dots,N-2}^{j_k(p_\mathbf{n}+q_\mathbf{n})+p_\mathbf{n}}\ \sum_{i_{N-1}=j_{N-1}(p_\mathbf{n}+q_\mathbf{n})+p_\mathbf{n}+1}^{(j_{N-1}+1)(p_\mathbf{n}+q_\mathbf{n})}\ \sum_{i_N=j_N(p_\mathbf{n}+q_\mathbf{n})+p_\mathbf{n}+1}^{(j_N+1)(p_\mathbf{n}+q_\mathbf{n})}\Lambda_\mathbf{i},$$

and so on. The last two terms are

$$W(2^{N-1},\mathbf{n},x,\mathbf{j})=\sum_{i_k=j_k(p_\mathbf{n}+q_\mathbf{n})+p_\mathbf{n}+1,\ k=1,\dots,N-1}^{(j_k+1)(p_\mathbf{n}+q_\mathbf{n})}\ \sum_{i_N=j_N(p_\mathbf{n}+q_\mathbf{n})+1}^{j_N(p_\mathbf{n}+q_\mathbf{n})+p_\mathbf{n}}\Lambda_\mathbf{i},$$

$$W(2^{N},\mathbf{n},x,\mathbf{j})=\sum_{i_k=j_k(p_\mathbf{n}+q_\mathbf{n})+p_\mathbf{n}+1,\ k=1,\dots,N}^{(j_k+1)(p_\mathbf{n}+q_\mathbf{n})}\Lambda_\mathbf{i}$$

where $q_\mathbf{n}=o\left(\left[\widehat{\mathbf{n}}(h_H\phi_x(h_K))^{(1+2N)}\right]^{1/(2N)}\right)$ and $p_\mathbf{n}=\left[(\widehat{\mathbf{n}}h_H\phi_x(h_K))^{1/(2N)}/s_\mathbf{n}\right]$ with $s_\mathbf{n}=o\left(\left[\widehat{\mathbf{n}}(h_H\phi_x(h_K))^{(1+2N)}\right]^{1/(2N)}q_\mathbf{n}^{-1}\right)$. Because of (H11) all the sequences $q_\mathbf{n}$, $p_\mathbf{n}$ and $s_\mathbf{n}$ tend to infinity.
Now, Define for each $i=1,\dots,2^N$,

$$T(\mathbf{n},x,i)=\sum_{\mathbf{j}\in\mathcal{J}}W(i,\mathbf{n},x,\mathbf{j}).$$

where $\mathcal{J}=\{0,\dots,r_1-1\}\times\dots\times\{0,\dots,r_N-1\}$ with $r_k=n_k(p_\mathbf{n}+q_\mathbf{n})^{-N}$. So, we have

$$\left(\widehat{\mathbf{n}}(\sigma^2(x))\right)^{-1/2}S_\mathbf{n}$$

$$=\left[\sqrt{\widehat{\mathbf{n}}}\sigma(x)\right]^{-1}\left(T(\mathbf{n},x,1)+\sum_{i=2}^{2N}T(\mathbf{n},x,i)\right).$$

Thus, the proof of the asymptotic normality of $\left(\widehat{\mathbf{n}}(\sigma^2(x))\right)^{-1/2}S_\mathbf{n}$ is reduced to the proofs of the following results

$$(13)\qquad Q_1\equiv\left|E\left[\exp\left[iuT(\mathbf{n},x,1)\right]\right]-\prod_{\substack{j_k=0\\k=1,\dots,N}}^{r_k-1}E\left[\exp\left[iuW(1,\mathbf{n},x,\mathbf{j})\right]\right]\right|\to 0$$

(14) $$Q_2 \equiv \widehat{\mathbf{n}}^{-1} E\Big[\sum_{i=2}^{2^N} T(\mathbf{n}, x, i)\Big]^2 \to 0$$

(15) $$Q_3 \equiv \widehat{\mathbf{n}}^{-1} \sum_{\mathbf{j}\in\mathcal{J}} E\Big[W(1, \mathbf{n}, x, \mathbf{j})\Big]^2 \to \sigma^2(x)$$

(16) $$Q_4 \equiv \widehat{\mathbf{n}}^{-1} \sum_{\mathbf{j}\in\mathcal{J}} E\Big[(W(1, \mathbf{n}, x, \mathbf{j}))^2 \mathbf{1}_{\{|W(1,\mathbf{n},x,\mathbf{j})|>\epsilon(\sigma^2(x)\widehat{\mathbf{n}})^{1/2}\}}\Big] \to 0 \text{ for all } \epsilon > 0.$$

For (11) : We numerate the $M = \prod_{k=1}^N r_k = \widehat{\mathbf{n}}(p_\mathbf{n} + q_\mathbf{n})^{-N} \leq \widehat{\mathbf{n}} p_\mathbf{n}^{-N}$ random variables $W(1, \mathbf{n}, x, \mathbf{j}); \mathbf{j} \in \mathcal{J}$ in the arbitrary way $\tilde{U}_1, ..., \tilde{U}_M$. For $\mathbf{j} \in \mathcal{J}$, let

$$I(1, \mathbf{n}, x, \mathbf{j}) = \{\mathbf{i} : j_k(p_\mathbf{n} + q_\mathbf{n}) + 1 \leq i_k \leq j_k(p_\mathbf{n} + q_\mathbf{n}) + p_\mathbf{n} \quad ; k = 1, ..N\}$$

then we have $\quad W(1, \mathbf{n}, x, \mathbf{j}) = \sum_{\mathbf{i}\in I(1,\mathbf{n},x,\mathbf{j})} \Lambda_\mathbf{i}$. Noting that each of the sets of site $I(1, \mathbf{n}, x, \mathbf{j})$ contains $p_\mathbf{n}^N$, these sets are distant of $p_\mathbf{n}$ at least. Further, we apply the lemma of Volkonski and Rozanov (1959) to the variable $\Big(\exp(iu\tilde{U}_1), ..., \exp(iu\tilde{U}_M)\Big)$. Since $\Big|\prod_{s=j+1}^{M} \exp[iu\tilde{U}_s]\Big| \leq 1$, then :

$$\begin{aligned}
Q_1 &= \Big|E[\exp\big[iuT(\mathbf{n}, x, 1)\big]] - \prod_{\substack{j_k=0\\k=1,\dots,N}}^{r_k-1} E[\exp\big[iuW(1, \mathbf{n}, x, \mathbf{j})\big]]\Big| \\
&= \left|E \prod_{\substack{j_k=0\\k=1,\dots,N}}^{r_k-1} \exp\left[iuW(1, \mathbf{n}, x, \mathbf{j})\right] - \prod_{\substack{j_k=0\\k=1,\dots,N}}^{r_k-1} E\exp\big[iuW(1, \mathbf{n}, x, \mathbf{j})\big]\right| \\
&\leq \sum_{k=1}^{M-1}\sum_{j=k+1}^{M} \Big|E\Big(\exp[iu\tilde{U}_k] - 1\Big)\Big(\exp[iu\tilde{U}_j] - 1\Big)\prod_{s=j+1}^{M}\exp[iu\tilde{U}_s] \\
&\qquad\qquad - E\Big(\exp[iu\tilde{U}_k] - 1\Big)E\Big(\exp[iu\tilde{U}_j] - 1\Big)\prod_{s=j+1}^{M}\exp[iu\tilde{U}_s]\Big| \\
&= \sum_{k=1}^{M-1}\sum_{j=k+1}^{M} \Big|E\Big(\exp[iu\tilde{U}_k] - 1\Big)\Big(\exp[iu\tilde{U}_j] - 1\Big) - E\Big(\exp[iu\tilde{U}_k] - 1\Big)E\Big(\exp[iu\tilde{U}_j] - 1\Big)\Big| \\
&\qquad\qquad \times\Big|\prod_{s=j+1}^{M}\exp[iu\tilde{U}_s]\Big| \\
&\leq \sum_{k=1}^{M-1}\sum_{j=k+1}^{M} \Big|E\Big(\exp[iu\tilde{U}_k] - 1\Big)\Big(\exp[iu\tilde{U}_j] - 1\Big) - E\Big(\exp[iu\tilde{U}_k] - 1\Big)E\Big(\exp[iu\tilde{U}_j] - 1\Big)\Big|.
\end{aligned}$$

Let $\tilde{I}_j$ be the set of sites among the $I(1,\mathbf{n},x,\mathbf{j})$ such that $\tilde{U}_j = \sum_{\mathbf{i}\in\tilde{I}_j} D_{\mathbf{i}}$. The lemma of Carbon et al. (1997) and assumption (3), lead :

$$\left|E\left(\exp[iu\tilde{U}_k]-1\right)\left(\exp[iu\tilde{U}_j]-1\right)-E\left(\exp[iu\tilde{U}_k]-1\right)E\left(\exp[iu\tilde{U}_j]-1\right)\right| \leq C\varphi\left(d(\tilde{I}_j,\tilde{I}_k)\right)p_{\mathbf{n}}^N.$$

So

$$\begin{aligned}
Q_1 &\leq Cp_{\mathbf{n}}^N \sum_{k=1}^{M-1}\sum_{j=k+1}^{M} \varphi\left(d(\tilde{I}_j,\tilde{I}_k)\right) \\
&\leq Cp_{\mathbf{n}}^N M \sum_{k=2}^{M} \varphi\left(d(\tilde{I}_1,\tilde{I}_k)\right) \\
&\leq Cp_{\mathbf{n}}^N M \sum_{i=1}^{\infty} \sum_{k:iq_{\mathbf{n}}\leq d(\tilde{I}_1,\tilde{I}_k)<(i+1)q_{\mathbf{n}}} \varphi\left(d(\tilde{I}_1,\tilde{I}_k)\right) \\
&\leq Cp_{\mathbf{n}}^N M \sum_{i=1}^{\infty} i^{N-1}\varphi(iq_{\mathbf{n}}) \\
&\leq C\widehat{\mathbf{n}}q_{\mathbf{n}}^{-\delta} \sum_{i=1}^{\infty} i^{N-1-\delta},
\end{aligned}$$

by (6). This last tends to zero by the fact that $\widehat{\mathbf{n}}q_{\mathbf{n}}^{-\delta} \to 0$ (see (H11)).

Proof of (12) : We have

$$Q_2 \equiv \widehat{\mathbf{n}}^{-1}E\Big[\sum_{i=2}^{2^N} T(\mathbf{n},x,i)\Big]^2 = \widehat{\mathbf{n}}^{-1}\left(\sum_{i=2}^{2^N} E\left[T(\mathbf{n},x,i)\right]^2 + \sum_{\substack{i,j=2,\dots,2^N \\ i\neq j}} E\left[T(\mathbf{n},x,i)\right]\left[T(\mathbf{n},x,j)\right]\right).$$

By Cauchy-Schwartz inequality, we get :

$$\forall\ 2\leq i\leq 2^N : \widehat{\mathbf{n}}^{-1}E\left[T(\mathbf{n},x,i)\right]\left[T(\mathbf{n},x,j)\right] \leq \left(\widehat{\mathbf{n}}^{-1}E\left[T(\mathbf{n},x,i)\right]^2\right)^{1/2}\left(\widehat{\mathbf{n}}^{-1}E\left[T(\mathbf{n},x,j)\right]^2\right)^{1/2}.$$

Then, it suffices to prove that

$$\widehat{\mathbf{n}}^{-1}E\left[T(\mathbf{n},x,i)\right]^2 \to 0 \quad ;\forall\ 2\leq i\leq 2^N.$$

We will prove this for $i=2$, the case where $i\neq 2$ is similar. We have $T(\mathbf{n},x,2) = \sum_{\mathbf{j}\in\mathcal{J}} W(2,\mathbf{n},x,\mathbf{j}) = \sum_{j=1}^{M}\hat{U}_j$, where we enumerate the $W(2,\mathbf{n},x,\mathbf{j})$ in the arbitrary way $\hat{U}_1,...,\hat{U}_M$.

Then :

$$\begin{aligned}E\left[T(\mathbf{n},x,2)\right]^2 &= \sum_{i=1}^{M} Var\left(\hat{U}_i\right) + \sum_{i=1}^{M}\sum_{\substack{j=1\\ i\neq j}}^{M} Cov\left(\hat{U}_i,\hat{U}_j\right)\\ &= A_1 + A_2.\end{aligned}$$

The stationarity of the process $(X_{\mathbf{i}}, Y_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^N}$, implies that :

$$\begin{aligned}Var\left(\hat{U}_i\right) &= Var\left(\sum_{\substack{i_k=1\\ k=1,\ldots,N-1}}^{p_{\mathbf{n}}}\sum_{i_N=1}^{q_{\mathbf{n}}}\Lambda_{\mathbf{i}}\right)^2\\ &= p_{\mathbf{n}}^{N-1}q_{\mathbf{n}}\; Var\left[\Lambda_{\mathbf{i}}\right]\\ &\quad + \sum_{\substack{i_k=1\\ k=1,\ldots,N-1}}^{p_{\mathbf{n}}}\sum_{i_N=1}^{q_{\mathbf{n}}}\underset{\mathbf{i}\neq\mathbf{j}}{\sum_{\substack{j_k=1\\ k=1,\ldots,N-1}}^{p_{\mathbf{n}}}}\sum_{j_N=1}^{q_{\mathbf{n}}} E[\Lambda_{\mathbf{i}}\Lambda_{\mathbf{j}}].\end{aligned}$$

We proved above that $Var\left[\Lambda_{\mathbf{i}}\right] < C$ (see 11). By Lemma 4.5.2, we have :

$$|E[\Lambda_{\mathbf{i}}\Lambda_{\mathbf{j}}]| \leq C(h_H\phi_x(h_K))^{-1}\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right). \tag{17}$$

Then, we deduce that

$$\begin{aligned}Var\left[\hat{U}_i\right] &\leq Cp_{\mathbf{n}}^{N-1}q_{\mathbf{n}}\left(1 + (h_H\phi_x(h_K))^{-1}\sum_{\substack{i_k=1\\ k=1,\ldots,N-1}}^{p_{\mathbf{n}}}\sum_{i_N=1}^{q_{\mathbf{n}}}\left(\varphi\left(\|\mathbf{i}\|\right)\right)\right)\\ &\leq Cp_{\mathbf{n}}^{N-1}q_{\mathbf{n}}(h_H\phi_x(h_K))^{-1}\sum_{\substack{i_k=1\\ k=1,\ldots,N-1}}^{p_{\mathbf{n}}}\sum_{i_N=1}^{q_{\mathbf{n}}}\left(\varphi(\|\mathbf{i}\|)\right).\end{aligned}$$

Consequently, we have :

$$A_1 \leq CMp_{\mathbf{n}}^{N-1}q_{\mathbf{n}}(h_H\phi_x(h_K))^{-1}\sum_{i=1}^{\infty} i^{N-1}\left(\varphi(i)\right).$$

Let

$$\begin{aligned}I\left(2,\mathbf{n},x,\mathbf{j}\right) &= \Big\{\mathbf{i} : j_k(p_{\mathbf{n}}+q_{\mathbf{n}})+1 \leq i_k \leq j_k(p_{\mathbf{n}}+q_{\mathbf{n}})+p_{\mathbf{n}},\quad 1\leq k\leq N-1;\\ &\qquad +j_N(p_{\mathbf{n}}+q_{\mathbf{n}})+p_{\mathbf{n}}+1 \leq i_N \leq (j_N+1)(p_{\mathbf{n}}+q_{\mathbf{n}})\Big\}.\end{aligned}$$

The variable $U\left(2,\mathbf{n},x,\mathbf{j}\right)$ is the sum of the $\Lambda_{\mathbf{i}}$ such that $\mathbf{i}$ is in $I\left(2,\mathbf{n},x,\mathbf{j}\right)$. Since $p_{\mathbf{n}}>q_{\mathbf{n}}$, if $\mathbf{i}$ and $\mathbf{i}'$ are respectively in the two different sets $I\left(2,\mathbf{n},x,\mathbf{j}\right)$ and $I\left(2,\mathbf{n},x,\mathbf{j}'\right)$; then $i_k\neq i'_k$ for un certain k such that $1\leq k\leq N$ and $\|\mathbf{i}-\mathbf{i}'\|>q_{\mathbf{n}}$.
By using the definition of A_2, the stationarity of the process and (15), we have :

$$A_2\leq \sum_{\substack{j_k=1\\k=1,\dots,N}}^{n_k}\ \sum_{\substack{i_k=1\\k=1,\dots,N\\\|\mathbf{i}-\mathbf{j}\|>q_{\mathbf{n}}}}^{n_k} E[\Lambda_{\mathbf{i}}\Lambda_{\mathbf{j}}]\leq C(h_H\phi_x(h_K))^{-1}\widehat{\mathbf{n}}\sum_{\substack{i_k=1\\k=1,\dots,N\\\|\mathbf{i}\|>q_{\mathbf{n}}}}^{n_k}\left(\varphi(\|\mathbf{i}\|)\right)$$

and

$$A_2\leq C(h_H\phi_x(h_K))^{-1}\widehat{\mathbf{n}}\sum_{i=q_{\mathbf{n}}}^{\infty} i^{N-1}\left(\varphi(i)\right).$$

We deduce that :

$$\begin{aligned}\widehat{\mathbf{n}}^{-1}E\left[T(\mathbf{n},x,2)\right]^2 &\leq CMp_{\mathbf{n}}^{N-1}q_{\mathbf{n}}\widehat{\mathbf{n}}^{-1}(h_H\phi_x(h_K))^{-1}\sum_{i=1}^{\infty}i^{N-1-\delta}\\ &\quad +C(h_H\phi_x(h_K))^{-1}\sum_{i=q_{\mathbf{n}}}^{\infty}i^{N-1-\delta}.\end{aligned}$$

From $(p_{\mathbf{n}}+q_{\mathbf{n}})^{-N}p_{\mathbf{n}}^{N-1}q_{\mathbf{n}}=(p_{\mathbf{n}}+q_{\mathbf{n}})^{-N}p_{\mathbf{n}}^{N}\left(\frac{q_{\mathbf{n}}}{p_{\mathbf{n}}}\right)\leq\frac{q_{\mathbf{n}}}{p_{\mathbf{n}}}$, we get :

$$\begin{aligned}CMp_{\mathbf{n}}^{N-1}q_{\mathbf{n}}\widehat{\mathbf{n}}^{-1}(h_H\phi_x(h_K))^{-1} &= \widehat{\mathbf{n}}(p_{\mathbf{n}}+q_{\mathbf{n}})^{-N}p_{\mathbf{n}}^{N-1}q_{\mathbf{n}}\widehat{\mathbf{n}}^{-1}(h_H\phi_x(h_K))^{-1}\\ &\leq \left(\frac{q_{\mathbf{n}}}{p_{\mathbf{n}}}\right)(h_H\phi_x(h_K))^{-1}\\ &= q_{\mathbf{n}}s_{\mathbf{n}}\left(\widehat{\mathbf{n}}(h_H\phi_x(h_K))\right)^{\frac{-1}{2N}}(h_H\phi_x(h_K))^{-1}\\ &= q_{\mathbf{n}}s_{\mathbf{n}}\left(\widehat{\mathbf{n}}(h_H\phi_x(h_K))^{(1+2N)}\right)^{\frac{-1}{2N}}.\end{aligned}$$

By the definitions of $q_{\mathbf{n}}$ and $s_{\mathbf{n}}$, this last term converges to $\to 0$. Moreover, we have :

$$C(h_H\phi_x(h_K))^{-1}\sum_{i=q_{\mathbf{n}}}^{\infty}i^{N-1-\delta}\leq C(h_H\phi_x(h_K))^{-1}\int_{q_{\mathbf{n}}}^{\infty}t^{N-1-\delta}dt=C(h_H\phi_x(h_K))^{-1}q_{\mathbf{n}}^{N-\delta}.$$

This last term converges to zero by (H11) and ends the proof of (12).
For (13) : Let us use the following decomposition of small and big blocks

$$S'_{\mathbf{n}}=T\left(\mathbf{n},x,1\right)\qquad\qquad S''_{\mathbf{n}}=\sum_{i=2}^{2^N}T\left(\mathbf{n},x,i\right).$$

Then, we can write :

$$\widehat{\mathbf{n}}^{-1}E\left[S'_{\mathbf{n}}\right]^2=\widehat{\mathbf{n}}^{-1}E[S_{\mathbf{n}}^2]+\widehat{\mathbf{n}}^{-1}E\left[S''_{\mathbf{n}}\right]^2-2\widehat{\mathbf{n}}^{-1}E[S_{\mathbf{n}}S''_{\mathbf{n}}].$$

Lemma 4.4.7 and (12) imply respectively that $\widehat{\mathbf{n}}^{-1}E\left(S_{\mathbf{n}}\right)^{2}=\widehat{\mathbf{n}}^{-1}Var\left(S_{\mathbf{n}}\right)\rightarrow\sigma^{2}(x)$ and $\widehat{\mathbf{n}}^{-1}E\left[S_{\mathbf{n}}''\right]^{2}\rightarrow 0$.
Then, to show that $\widehat{\mathbf{n}}^{-1}E\left[S_{\mathbf{n}}'\right]^{2}\rightarrow\sigma^{2}(x)$, it suffices to remark that $\widehat{\mathbf{n}}^{-1}E[S_{\mathbf{n}}S_{\mathbf{n}}'']\rightarrow 0$ because, by Cauchy-Schwartz's inequality, we can write :

$$\left|\widehat{\mathbf{n}}^{-1}E[S_{\mathbf{n}}S_{\mathbf{n}}'']\right|\leq\widehat{\mathbf{n}}^{-1}E\left|S_{\mathbf{n}}S_{\mathbf{n}}''\right|\leq\left(\widehat{\mathbf{n}}^{-1}E[S_{\mathbf{n}}^{2}]\right)^{1/2}\left(\widehat{\mathbf{n}}^{-1}E[S_{\mathbf{n}}''^{2}]\right)^{1/2}.$$

Recall that $T\left(\mathbf{n},x,1\right)=\sum_{\mathbf{j}\in\mathcal{J}}W\left(1,\mathbf{n},x,\mathbf{j}\right)$, so

$$\begin{aligned}\widehat{\mathbf{n}}^{-1}E\left(S_{\mathbf{n}}'\right)^{2} &= \widehat{\mathbf{n}}^{-1}\sum_{\substack{j_k=0\\ k=1,\ldots,N}}^{r_k-1}E\left[W\left(1,\mathbf{n},x,\mathbf{j}\right)\right]^{2}\\ &+ \widehat{\mathbf{n}}^{-1}\times\sum_{\substack{j_k=0\\ k=1,\ldots,N}}^{r_k-1}\ \sum_{\substack{i_k=0\\ k=1,\ldots,N\\ i_k\neq j_k \text{ for some k}}}^{r_k-1}Cov(W(1,\mathbf{n},x,\mathbf{j}),W(1,\mathbf{n},x,\mathbf{i}))\end{aligned}$$

By similar arguments used above for A_2, this last term is not greater than

$$\begin{aligned}&C(h_H\phi_x(h_K))^{-1}\sum_{\substack{i_k=1\\ k=1,\ldots,N\\ \|\mathbf{i}\|>q_{\mathbf{n}}}}^{r_k-1}(\varphi(\|\mathbf{i}\|))\\ &\leq C(h_H\phi_x(h_K))^{-1}\sum_{i=q_{\mathbf{n}}}^{\infty}i^{N-1}\left(\varphi(i)\right)\leq C(h_H\phi_x(h_K))^{-1}q_{\mathbf{n}}^{N-\delta}\rightarrow 0.\end{aligned}$$

So $Q_3\rightarrow\sigma^2(x)$. This ends the proof.

Proof of 14 : Since $|\Lambda_{\mathbf{i}}|\leq C(h_H\phi_x(h_K))^{-1/2}$, we have : $|W\left(1,\mathbf{n},x,\mathbf{j}\right)|\leq Cp_{\mathbf{n}}^{N}(h_H\phi_x(h_K))^{-1/2}$. Then we deduce that

$$Q_4\leq Cp_{\mathbf{n}}^{2N}(h_H\phi_x(h_K))^{-1}\widehat{\mathbf{n}}^{-1}\sum_{\substack{j_k=0\\ k=1,\ldots,N}}^{r_k-1}P\left[\left|W\left(1,\mathbf{n},x,\mathbf{j}\right)\right|>\epsilon\left(\sigma^2(x)\widehat{\mathbf{n}}\right)^{1/2}\right].$$

We have $\left|W\left(1,\mathbf{n},x,\mathbf{j}\right)\right|/\left(\left(\sigma^2(x)\widehat{\mathbf{n}}\right)^{1/2}\right)\leq Cp_{\mathbf{n}}^{N}\left(\widehat{\mathbf{n}}h_H\phi_x(h_K)\right)^{-1/2}=C\left(s_{\mathbf{n}}\right)^{-N}\rightarrow 0$, because $p_{\mathbf{n}}=\left[\left(\widehat{\mathbf{n}}h_H\phi_x(h_K)\right)^{1/(2N)}/s_{\mathbf{n}}\right]$ and $s_{\mathbf{n}}\rightarrow\infty$.
So, for all ϵ and $\mathbf{j}\in\mathcal{J}$; if $\widehat{\mathbf{n}}$ is great enough, then $P\left[W\left(1,\mathbf{n},x,\mathbf{j}\right)>\epsilon\left(\sigma^2(x)\widehat{\mathbf{n}}\right)^{1/2}\right]=0$. Then $Q_4=0$ for $\widehat{\mathbf{n}}$ great enough. This yields the proof.

References

Anselin,L. ; Florax,R.J.G.M. ; Rey, S.J. (2004) Econometrics for spatial model : recent advances. *Advances in spatial econometrics*, 1-25, Springer, Berlin,

Biau, G., Cadre, B. (2004) Nonparametric spatial prediction. *Statist. Inference Stoch. Processes*, 7, 327-349.

Bosq, D. (2000) *Linear Processes in Function Spaces : Theory and Applications.* Lecture Notes in Statistics, 149. Springer. New-York.

Cressie, N.A. (1993) *Statistics for spatial data.* Wiley, New York.

Carbon, M., Hallin, M., Tran, L.T. (1996) Kernel density estimation for random fields : the L^1 -theory. *J. Nonparametric Statist.*, 6, 157-170.

Carbon, M., Tran, L. T., Wu, B. (1997) Kernel density estimation for random fields. *Statist. Probab. Lett.*, 36, 115-125.

Carbon, M., Francq, C., Tran, L.T. (2007) Kernel regression estimation for random fields. *J. Statist. Plann. Inference*, 137, 778-798.

Dabo-Niang, S. and Laksaci, A. (2010). Note on conditional mode estimation for functional dependent data, *Statistica*, **70**, 83-94.

Dabo-Niang, S. and Thiam, B. (2010). Robust quantile estimation and prediction for spatial processes. *Statist. Probab. Lett.* **80**, 1447-1458.

Dabo-Niang, S., Rachdi, M. and Yao, A-F. (2011a). Spatial kernel regression estimation and prediction for functional random fields. *Far. East. J. Statist.*, **37**, 77-113.

Dabo-Niang,S. Kaid,Z. Laksaci, A. (2011b) Sur la rï£¡gression quantile pour variable explicative fonctionnelle : Cas des donnï£¡es spatiales *C. R., Math., Acad. Sci. Paris.* **349,** 1287-1291.

Diggle, P., Ribeiro, P.J. (2007) *Model-Based Geostatistics.* Springer, New York.

Ezzahrioui, M. and Ould-Saï£¡d, E. (2008). Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data. *J. Nonparametr. Stat.* **20**, 3-18.

Ferraty, F.,Vieu, P. (2006) *Nonparametric Functional Data Analysis. Theory and Practice.* Springer-Verlag, New York.

Ferraty, F., Laksaci, A., Tadj, A. and Vieu, P. (2010). Rate of uniform consistency for nonparametric estimates with functional variables. *J. Statist. Plann. Inference.* **140**,335-352.

Gao, J. ; Lu, Z. ; Tjï£¡stheim, D. (2008) Moment inequalities for spatial processes. *Statist. Probab. Lett.* **78** 687–697.

Giraldo, R.(2009) Geostatistical Analysis of Functional Data. Ph.D. thesis. Universitat Politï£¡cnica de Catalunya.

Guyon, X. (1987) Estimation d'un champ par pseudo-vraisemblance conditionnelle : Etude asymptotique et application au cas Markovien. *Proceedings of the Sixth Franco-Belgian Meeting of Statisticians.*

Guyon, X. (1995) *Random fields on a network. Modeling, statistics and applications. Probability and its applications.* Springer-Verlag, New York.

Hallin, M., Lu, Z., Tran, L.T. (2004) Local linear spatial regression. *Ann. Statist.*, 32, 2469-2500.

Laksaci, A. and Maref, F. (2009). Estimation non paramï£¡trique de quantiles conditionnels pour des variables fonctionnelles spatialement dï£¡pendantes. *C. R., Math., Acad. Sci. Paris.* **347**, 1075-1080.

Laksaci, A., Madani, F. and Rachdi, M. (2011). Kernel conditional density estimation when the regressor is valued in a semi-metric space *Commun. Stat., Theory and Methods,* to appear.

Li, J., Tran, L.T. (2009) Nonparametric estimation of conditional expectation. *J. Statist. Plann. Inference*, **139**, 164-175.

Lu, Z., Chen, X. (2004) Spatial kernel regression : weak consistency. *Statist. Probab. Lett.*, 68, 125-136.

Ould Abdi, A.; Diop, A.; Dabo-Niang, S.; Ould Abdi, S.A. (2010b) Estimation non paramï£¡trique du mode conditionnel dans le cas spatial. *C. R. Math. Acad. Sci. Paris* **348**, 815–819.

Ramsay, J.O., Silverman, B.W. (2002) *Applied Functional Data Analysis.* Springer, New York.

Ramsay, J.O., Silverman, B.W. (2005) *Functional Data Analysis, Second Edition.* Springer, New York.

Ramsay, J. (2008). Fda problems that I like to talk about. Personal communication.

Ripley, B. (1981). *Spatial Statistics.* Wiley, New York.

Chapitre 4

La normalité asymptotique et la convergence en norme L^p de l'estimateur spatial des quantiles conditionnels

L'objet de ce chapitre est l'estimation non paramétrique spatiale des quantiles conditionnels. Plus précisément, on estime ce modèle par la méthode du noyau en utilisant l'inverse de l'estimateur à double noyaux de la fonction de répartition conditionnelle. Sous des conditions assez générales, on démontre la convergence en norme L^p et la normalité asymptotique de cet estimateur. De même que pour les autres chapitres, le défi principal est de contourner le problème de l'absence d'une mesure de référence, l'absence d'une structure banachique, et d'une relation d'ordre sur les indices spatiaux,.... Notons que ces derniers sont primordiaux pour utiliser les outils classiques tels que l'intégration et la dérivabilité.*Ce Chapitre fait l'objet d'un article dans les Comptes rendus de l'académie des sciences de Paris : C.R.A.S.*

SPATIAL CONDITIONAL QUANTILE REGRESSION: WEAK CONSISTENCY OF A KERNEL ESTIMATE

SOPHIE DABO-NIANG [♯], ZOULIKHA KAID[§] AND ALI LAKSACI[§]

Abstract. We consider a conditional quantile regression model for spatial data. More precisely, given a strictly stationary random field $Z_{\mathbf{i}} = (X_{\mathbf{i}},\ Y_{\mathbf{i}})_{\mathbf{i}\in\mathbb{N}^N}$, we investigate a kernel estimate of the conditional quantile regression function of the univariate response variable $Y_{\mathbf{i}}$ given the functional variable $X_{\mathbf{i}}$. The principal aim of the paper is to prove the convergence (with rate) in L^p norm and the asymptotic normality of the estimator.

Key Words: Kernel conditional quantile estimation; Kernel regression estimation; Spatial process.

♯ Laboratory EQUIPPE, Université Charles De Gaulle, Lille 3, maison de la recherche, domaine du pont de bois, BP 60149, 59653 Villeneuve d'ascq cedex, France. sophie.dabo@univ-lille3.fr.

§ Département de Mathématiques, Univ. Djillali Liabès, BP 89, 22000 Sidi Bel Abbès, Algérie. kadezolekha@yahoo.com, alilak@yahoo.fr.

4.1 Introduction

Conditional quantile estimation is an important field in statistics which dates back to Stone (1977) and has been widely studied in the non-spatial case. It is useful in all domain of statistics, such as time series, survival analysis and growth charts among others, see Koenker (2000, 2005) for a review. There exist an extensive literature and various nonparametric approaches in conditional quantile estimation in the non spatial case for independent samples and dependent non-functional or functional observations. Among the lot of papers dealing with conditional quantile estimation in finite dimension, one can refer for example to key works of Stute (1986), Samanta (1989) and Berlinet *et al.* (1989), Portnoy (1991), Koul and Mukherjee (1994), Honda (2000), Gannoun *et al.* (2003), Yu *et al.* (2003).

Potential applications of quantile regression to spatial data are without number. Indeed, there is an increasing number of situations coming from different fields of applied sciences (soil science, geology, oceanography, econometrics, epidemiology, environmental science, forestry,...), where the influence of a vector of covariates on some response variable is to be studied in a context of spatial dependence. The literature on spatial models is relatively abundant, see for example, Guyon (1995), Anselin and Florax (1995), Cressie (1991) or Ripley (1981) for a list of references.

In our knowledge, only the papers of Koencker and Mizera (2004), Hallin *et al.* (2009), Abdi *et al.* (2010a, 2010b), Dabo-Niang and Thiam (2010) have paid attention to study nonparametric quantile regression for finite dimensional random fields while Laksaci and Maref (2010) have considered infinite dimensional fields. These last work deals with almost sure consistency of the conditional consistency of a kernel conditional quantile estimate. The work of Hallin *et al.* (2010) deals with local linear spatial conditional quantile regression estimation. The method of Koencker and Mizera (2004) is a spatial smoothing technique rather than a spatial (auto)regression one and do not take into account the spatial dependency structure of the data. The results of Abdi *et al.* (2010a, 2010b) concerned respectively consistency in $p-$mean ($p > 1$) and asymptotic normality and of a kernel estimate of the conditional regression function for spatial processes. Dabo-Niang and Thiam (2010) considered the L_1 consistency of the local linear and double kernel conditional quantile estimate.

As in the non-spatial case, conditional quantile estimation is useful for some non-parametric prediction models and is used as an alternative to classical spatial regression estimation models for non-functional data (see Biau and Cadre (2004), Lu and Chen (2002, 2004), Hallin, Lu and Tran (2004a), Dabo-Niang and Yao (2007)). Spatial conditional quantile is of wide interest in the modeling of spatial dependence and in the construction of confidence (predictive) intervals. The purpose of this paper is to estimate the conditional quantile regression for spatial functional data.

Recall that a recent and restrictive attention has been paid to nonparametric estimation of the conditional quantile of a scalar variable Y given a functional variable ($X = X_t$, $t \in \mathbb{R}$) when observations are over an interval $T \subset \mathbb{R}$. The first results concerning the conditional quantile estimation adapted to non-spatial functional data were obtained by Cardot et *al.* (2004). They used the B-spline approach to study a linear model of regression on quantiles when the explanatory variable takes values in a Hilbert space and established the L^2-convergence rate of the estimate. In the nonparametric context, Ferraty *et al.* (2006) established the almost complete convergence of a kernel estimator of the conditional quantile when the observations are i.i.d. This last work has been extended to dependant case by Ferraty *et al.* (2005), with an application to climatologic data. Ezzahrioui and Ould-said (2008a, 2008b) have studied in both cases (i.i.d and strong mixing) the asymptotic normality of the kernel estimator when the explicative variable satisfies some general fractal condition. Recently, Dabo-Niang and Laksaci (2007) and Dabo-Niang and Laksaci (2009) stated asymptotic normality and convergence in L^p norm (in the i.i.d. and dependent cases) under less restrictive conditions closely related to the concentration properties on small balls probability of the underlying explanatory variable.

The main aim of this paper is to extend some of the results on quantile regression to the case of functional spatial processes. In our knowledge, this work is the first contribution on spatial quantile regression estimation for functional variables. Noting that, extending classical nonparametric conditional quantile estimation for dependent functional random variables to quantile regression for functional random fields is far from being trivial. This is due to the absence of any canonical ordering in the space, and of obvious definition of tail sigma-fields.

The paper is organized as follows. In Section 2 we provide the notations and the kernel quantile estimates. Section 3 is devoted to assumptions. Section 4 is devoted to the L_p convergence and the asymptotic normality results of the kernel quantile regression estimate, under mixing assumptions. Proofs and technical lemmas are given in Section 5.

4.2 The model

Consider $Z_{\mathbf{i}} = (X_{\mathbf{i}}, Y_{\mathbf{i}})$, $\mathbf{i} \in \mathbb{N}^N$ be a $\mathcal{F} \times \mathbb{R}$-valued measurable strictly stationary spatial process, defined on a probability space $(\Omega,\ \mathcal{A}, \mathbf{P})$, where $(\mathcal{F},\ d)$ is a semi-metric space. Let d denotes the semi-metric and $N \geq 1$. A point $\mathbf{i} = (i_1, ..., i_N) \in \mathbb{N}^N$ will be referred to as a site. We assume that the process under study $(Z_{\mathbf{i}})$ is observed over a rectangular domain $I_{\mathbf{n}} = \{\mathbf{i} = (i_1, ..., i_N) \in \mathbb{Z}^N\ ,\ 1 \leq i_k \leq n_k,\ k = 1, ..., N\}$, $\mathbf{n} = (n_1, ..., n_N) \in \mathbb{N}^N$. A point $\mathbf{i}$ will be referred to as a *site*. We will write $\mathbf{n} \to \infty$ if $\min\{n_k\} \to \infty$ and $|\frac{n_j}{n_k}| < C$ for a constant C such that $0 < C < \infty$ for all $j,\ k$ such that $1 \leq j, k \leq N$. For $\mathbf{n} = (n_1, ..., n_N) \in \mathbb{N}^N$, we set $\widehat{\mathbf{n}} = n_1 \times \cdots \times n_N$.

We assume that the $Z_{\mathbf{i}}$'s have the same distribution as (X, Y) and the regular version of the conditional probability of Y given X exists and admits a bounded probability density. For all $x \in \mathcal{F}$, we denote respectively by F^x and f^x the conditional distribution function and density of Y given $X = x$.

Let $\alpha \in]0, 1[$, the α^{th} conditional quantile noted $q_\alpha(x)$ is defined by

$$F^x(q_\alpha(x)) = \alpha.$$

To insure existence and unicity of $q_\alpha(x)$, we assume that F^x is strictly increasing. This last is estimated by

$$(1) \qquad \widehat{F}^x_{\mathbf{n}}(y) = \begin{cases} \dfrac{\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K_1\left(\frac{d(x,X_{\mathbf{i}})}{a_{\mathbf{n}}}\right) K_2\left(\frac{y-Y_{\mathbf{i}}}{b_{\mathbf{n}}}\right)}{\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K_1\left(\frac{d(x,X_{\mathbf{i}})}{a_{\mathbf{n}}}\right)} & \text{if} \quad \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K_1\left(\frac{d(x,X_{\mathbf{i}})}{a_{\mathbf{n}}}\right) \neq 0 \\ 0 & \text{else} \end{cases}.$$

where K_1 is a kernel, K_2 is a distribution function, $a_{\mathbf{n}}$ (resp $b_{\mathbf{n}}$) is a sequence of real numbers which converges to 0 when $\mathbf{n} \to \infty$.
The kernel estimate $\hat{q}_\alpha(x)$ of the conditional quantile $q_\alpha(x)$ is defined by

$$\widehat{F}^x(\hat{q}_\alpha(x)) = \alpha.$$

In the following, we fix a point x in $\mathcal{F}$ such that

$$P(X \in B(x,r)) = \phi_x(r) > 0,$$

where $B(x,h) = \{x' \in \mathcal{F}/ \ d(x',x) < h\}$.

4.3 Hypotheses

Throughout the paper, when no confusion will be possible, we will denote by C and C' any generic positive constant, and we denote by $g^{(j)}$ the derivative of order j of a function g. We will use the following hypotheses :
H_1 : F^x is of class $\mathcal{C}^1$ and $f^x(q_\alpha(x)) > 0$.
H_2 :$\exists \delta_1 > 0$, $\forall\,(y_1,y_2) \in [q_\alpha(x)-\delta_1, q_\alpha(x)+\delta_1]^2$, $\forall\,(x_1,x_2) \in N_x \times N_x$,

$$\left|F^{x_1}(y_1) - F^{x_2}(y_2)\right| \leq C\left(d^{b_1}(x_1,x_2) + |y_1-y_2|^{b_2}\right), \quad b_1 > 0, b_2 > 0.$$

where N_x is a small enough neighborhood of x
H_3 : There exist C_1 and C_2, $0 < C_1 < C_2 < \infty$ such that $C_1 \mathbb{I}_{[0,1]}(t) < K_1(t) < C_2 \mathbb{I}_{[0,1]}(t)$.
H_4 : K_2 is of class $\mathcal{C}^1$, of bounded derivative that verifies

$$\int_{\mathbb{R}} |t|^{b_2} K_2^{(1)}(t)dt < \infty.$$

Dependency conditions

In spatial dependent data analysis, the dependence of the observations has to be measured. Here we will consider the following two dependence measures :

Local dependence condition

In order to establish the same convergence rate as in the i.i.d. case (see Dabo-Niang and Laksaci (2007)), we need the following local dependency condition :
(2)

$$\begin{cases} \text{(i) For all } 1 \leq \mathbf{i}_1 < \dots \mathbf{i}_k \leq n, \ \text{ the conditional density of } \ (Y_{\mathbf{i}_1}, \dots Y_{\mathbf{i}_k}) \ \text{ given } \ (X_{\mathbf{i}_1}, \dots X_{\mathbf{i}_k}) \\ \text{exists and is bounded} \\ \text{(ii) For all } k \geq 2, \\ \text{we suppose that : there exists an increasing sequence } \ 0 < (v_k) < k, \ \text{ such that} \\ \max\left(\max_{\mathbf{i}_1 \dots \mathbf{i}_k \in \mathcal{I}_\mathbf{n}} P\left(d(X_{\mathbf{i}_j}, x) \leq r, 1 \leq j \leq k\right), \phi_x^k(r)\right) = O\left(\phi_x^{1+v_k}(r)\right). \end{cases}$$

Mixing condition

The spatial dependence of the process will be measured by means of $\alpha - mixing$. Then, we consider the $\alpha - mixing$ coefficients of the field $(Z_\mathbf{i}, \mathbf{i} \in \mathbb{N}^N)$, defined by : there exists a function $\varphi(t) \downarrow 0$ as $t \to \infty$, such that whenever $E,\ E^{'}$ subsets of $\mathbb{N}^N$ with finite cardinals,

$$\alpha\left(\mathcal{B}(E), \mathcal{B}\left(E^{'}\right)\right) = \sup_{B \in \mathcal{B}(E),\, C \in \mathcal{B}\left(E^{'}\right)} \left|\mathbf{P}(B \cap C) - \mathbf{P}(B)\mathbf{P}(C)\right|$$

$$\leq \psi\left(\mathrm{Card}(E), \mathrm{Card}\left(E^{'}\right)\right) \varphi\left(\mathrm{dist}\left(E, E^{'}\right)\right), \tag{3}$$

where $\mathcal{B}(E)$(*resp.* $\mathcal{B}\left(E^{'}\right)$) denotes the Borel σ-field generated by $(X_\mathbf{i},\ \mathbf{i} \in E)$ (*resp.* $\left(X_\mathbf{i},\ \mathbf{i} \in E^{'}\right)$), $\mathrm{Card}(E)$ (*resp.* $\mathrm{Card}\left(E^{'}\right)$) the cardinality of E (*resp.* $E^{'}$), $\mathrm{dist}\left(E, E^{'}\right)$ the Euclidean distance between E and $E^{'}$ and $\psi : \mathbb{N}^2 \to \mathbb{R}^+$ is a symmetric positive function nondecreasing in each variable.
Throughout the paper, it will be assumed that ψ satisfies either

$$\psi(n,m) \leq C \min(n,m), \qquad \forall n,m \in \mathbb{N} \tag{4}$$

or

$$\psi(n,m) \leq C(n+m+1)^{\widetilde{\beta}}, \qquad \forall n,m \in \mathbb{N} \tag{5}$$

for some $\widetilde{\beta} \geq 1$ and some $C > 0$. In the following, we will only considered Condition (4), one can extend esaily the asymptotic results proved here in the case of (5).
We assume also that the process satisfies the following mixing condition : a polynomial mixing condition :

$$\sum_{i=1}^{\infty} i^\delta \varphi(i) < \infty, \ \ \delta > N(p+2), \ \ p \geq 1. \tag{6}$$

If $N = 1$, then $X_\mathbf{i}$ is called strongly mixing. Many stochastic processes, among them various useful time series models satisfy strong mixing properties, which are relatively easy to check.

Conditions (4)-(5) are used in Tran (1990), Carbon *et al* (1996-1997), and are satisfied by many spatial models (see Guyon (1987) for some examples). In addition, we assume that

H_5 : $\exists 0 < \tau < (\delta - 5N)/2N$, η_0, $\eta_1 > 0$, such that $\widehat{\mathbf{n}}^\tau b_{\mathbf{n}} \to \infty$ and

$$C\widehat{\mathbf{n}}^{\frac{(5+2\tau)N-\delta}{\delta}+\eta_0} \leq \phi_x(a_{\mathbf{n}}),$$

where δ is introduced in (6).

Remark 2 If (6) is satisfied, then $\varphi(i) \leq Ci^{-\delta}$.

4.4 Main results

4.4.1 Weak consistency

This section contains results on pointwise consistency in p-mean. Let x be fixed, we give a rate of convergence of $\hat{q}_\alpha(x)$ to $q_\alpha(x)$ under some general conditions.

Theorem 4.4.1 *Under the hypotheses $H_1 - H_5$, (4), then, for all $p \geq 1$, we have*

$$\left\|\hat{q}_\alpha(x) - q_\alpha(x)\right\|_p = \left(E\left|\hat{q}_\alpha(x) - q_\alpha(x)\right|^p\right)^{1/p} = O\left((a_{\mathbf{n}})^{b_1} + (b_{\mathbf{n}})^{b_2}\right) + O\left(\left(\frac{1}{\widehat{\mathbf{n}}\,\phi_x(a_{\mathbf{n}})}\right)^{\frac{1}{2}}\right).$$

Proof
Let

$$K_{\mathbf{i}} = K_1\left(\frac{d(x, X_{\mathbf{i}})}{a_{\mathbf{n}}}\right),\ H_{\mathbf{i}}(y) = K_2\left(\frac{y - Y_{\mathbf{i}}}{b_{\mathbf{n}}}\right),\ W_{\mathbf{ni}} = W_{\mathbf{ni}}(x) = \frac{K_{\mathbf{i}}}{\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K_{\mathbf{i}}},$$

$$\hat{F}_N^x(y) = \frac{1}{\widehat{\mathbf{n}}EK_{\mathbf{1}}}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K_{\mathbf{i}}H_{\mathbf{i}}(y), \qquad \hat{F}_D^x = \frac{1}{\widehat{\mathbf{n}}EK_{\mathbf{1}}}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K_{\mathbf{i}}.$$

By hypothesis H_4, $\hat{F}_N^x(y)$ is of class $\mathcal{C}^1$, then we can write the following Taylor development :

$$\hat{F}_N^x(\hat{q}_\alpha(x)) = \hat{F}_N^x(q_\alpha(x)) + \hat{F}_N^{x(1)}(q_\alpha^*(x))\left(\hat{q}_\alpha(x) - q_\alpha(x)\right)$$

where $q_\alpha^*(x)$ is in the interval of extremities $q_\alpha(x)$ and $\hat{q}_\alpha(x)$. Thus,

$$\begin{aligned}
\hat{q}_\alpha(x) - q_\alpha(x) &= \frac{1}{\hat{F}_N^{x(1)}(q_\alpha^*(x))}\left(\hat{F}_N^x(\hat{q}_\alpha(x)) - \hat{F}_N^x(q_\alpha(x))\right)\\
&= \frac{1}{\hat{F}_N^{x(1)}(q_\alpha^*(x))}\left(\alpha\hat{F}_D^x - \hat{F}_N^x(q_\alpha(x))\right),
\end{aligned}$$

It is shown in Laksaci and Maref (2009) that under (H1)-(H5), (2) and (6) that

$$\hat{q}_\alpha(x) - q_\alpha(x) \to 0, \qquad \text{almost completely (a.co).}$$

So, by combining this consistency and the result of Lemma 11.17 in Ferraty and Vieu (2006, P.181) together with the fact that $q^*_\alpha(x)$ is lying between $\hat{q}_\alpha(x)$ and $q_\alpha(x)$, it follows that

$$\hat{F}_N^{x(1)}(q^*_\alpha(x)) - f^x(q_\alpha(x)) \to 0. \qquad a.co. \tag{7}$$

Since $f^x(q_\alpha(x)) > 0$, we can write

$$\exists C > 0 \quad \text{such that} \quad \left| \frac{1}{\hat{F}_N^{x(1)}(q^*_\alpha(x))} \right| \leq C \quad a.s.$$

It follows that

$$\|\hat{q}_\alpha(x) - q_\alpha(x)\|_p \ \leq C \left\| \alpha \hat{F}_D^x - \hat{F}_N^x(q_\alpha(x)) \right\|_p + \left(P\left(f_D^x = 0\right)\right)^{1/p}.$$

So, the rest of the proof is deduce from the following three lemmas.

Lemma 4.4.2 *Under $H_2 - H_4$, we have*

$$E\left[\alpha \widehat{F}_D^x - \widehat{F}_N^x(q_\alpha(x))\right] = O\left(a_{\mathbf{n}}^{b_1} + b_{\mathbf{n}}^{b_2}\right).$$

Lemma 4.4.3 *Under the hypotheses of Theorem 4.4.1, we have*

$$\left\| \alpha \widehat{F}_D^x - \widehat{F}_N^x(q_\alpha(x)) - E\left[\alpha \widehat{F}_D^x - \widehat{F}_N^x(q_\alpha(x))\right] \right\|_p = o\left(\left(\frac{1}{\widehat{\mathbf{n}}\, \phi_x(a_{\mathbf{n}})} \right)^{\frac{1}{2}} \right).$$

Lemma 4.4.4 *Under the hypotheses of Lemma 4.4.3, we have*

$$\left(P\left(\hat{F}_D^x = 0\right)\right)^{1/p} = o\left(\left(\frac{1}{\widehat{\mathbf{n}}\, \phi_x(a_{\mathbf{n}})} \right)^{\frac{1}{2}} \right).$$

4.4.2 Asymptotic normality

This section contains results on the asymptotic normality of the quantile estimator. For that we replace, respectively H_2 and H_4 by the following hypotheses.

H_2' : F^x satisfies H_2 and $\forall z \in N_x, F^z$ is of class $\mathcal{C}^1$ with respect to y, the conditional density f^x is such that $f^x(q_\alpha) > 0$ and $\forall (x_1, x_2) \in N_x \times N_x, \ \ \forall (y_1, y_2) \in \mathbb{R}^2$

$$|f^{x_1}(y_1) - f^{x_2}(y_2)| \leq \left(\|x_1 - x_2\|^{d_1} + |y_1 - y_2|^{d_2}\right), \quad d_1, d_2 > 0.$$

H_4' : K_2 satisfies H_4 and

$$\int |t|^{d_2} K_2^{(1)}(t)dt < \infty.$$

Theorem 4.4.5 *Under the hypotheses of Theorem 4.4.1 and H_2', H_4', (4) then, for any $x \in \mathcal{A}$, we have*

$$\left(\frac{(f^x(q_\alpha(x)))^2\widehat{\mathbf{n}}(\psi_{K_1}(a_{\mathbf{n}}))^2}{\psi_{K_1^2}(a_{\mathbf{n}})(\alpha(1-\alpha))}\right)^{(1/2)}(\widehat{q_\alpha}(x)-q_\alpha(x)-C_{\mathbf{n}}(x))\rightarrow_{n\rightarrow+\infty}N(0,1),$$

where

$$C_{\mathbf{n}}(x)=\frac{1}{f^x(q_\alpha(x))\psi_{K_1}(a_{\mathbf{n}})}\left(\alpha\psi_{K_1}(a_{\mathbf{n}})-E\left[K_1\left((a_{\mathbf{n}})^{-1}d(x,X)\right)F^X(q_\alpha(x))\right]\right)+O(b_{\mathbf{n}})$$

and

$$\mathcal{A}=\{x\in\mathcal{F},\ \frac{\psi_{K_1^2}(a_{\mathbf{n}})}{(\psi_{K_1}(a_{\mathbf{n}}))^2}\neq 0\}\quad with\ \psi_g(h)=-\int_0^1 g'(t)\phi_x(ht)dt.$$

Proof Firstly, observed that if H_5 is satisfied then, we have

$$\exists 0<\theta_1<1,\ \text{such that}\quad \widehat{\mathbf{n}}^{(-1+\theta_1)/(1+2N)}\leq\phi_x(a_{\mathbf{n}}). \tag{8}$$

Let

$$\Delta_{\mathbf{i}}=\frac{1}{\sqrt{EK_{\mathbf{1}}^2}}\left[\alpha K_{\mathbf{i}}-K_{\mathbf{i}}H_{\mathbf{i}}\left(q_\alpha\right)-E\left(\alpha K_{\mathbf{i}}-K_{\mathbf{i}}H_{\mathbf{i}}\left(q_\alpha\right)\right)\right].$$

By hypothesis H_4', $\hat{F}_N^x(y)$ is of class $\mathcal{C}^1$, then we can write the following Taylor development :

$$\hat{F}_N^x(\hat{q}_\alpha)=\hat{F}_N^x(q_\alpha)+\hat{F}_N^{x(1)}(q_\alpha^*)\left(\hat{q}_\alpha-q_\alpha\right)$$

where q_α^* is in the interval of extremities q_α and $\hat{q}_\alpha$. Thus,

$$\begin{aligned}\hat{q}_\alpha-q_\alpha &= \frac{1}{\hat{F}_N^{x(1)}(q_\alpha^*)}\left(\hat{F}_N^x(\hat{q}_\alpha)-\hat{F}_N^x(q_\alpha)\right)\\ &= \frac{1}{\hat{F}_N^{x(1)}(q_\alpha^*)}\left(\alpha\hat{F}_D^x-\hat{F}_N^x(q_\alpha)\right),\end{aligned}$$

$$\begin{aligned}&\left[\frac{f^x\left(q_\alpha\right)^2\widehat{\mathbf{n}}E^2K_{\mathbf{i}}}{\alpha(1-\alpha)EK_{\mathbf{i}}^2}\right]^{1/2}\left(\left[\hat{q}_\alpha-q_\alpha-C_{\mathbf{n}}(x)\right]\right)\\ &= \left[\frac{\widehat{\mathbf{n}}E^2K_{\mathbf{i}}}{\alpha(1-\alpha)EK_{\mathbf{i}}^2}\right]^{1/2}\left(\frac{f^x(q_\alpha)}{\hat{F}_N^{x(1)}(q_\alpha^*)}\left(\alpha\hat{F}_D^x-\hat{F}_N^x(q_\alpha)\right)-E\left(\alpha\hat{F}_D^x-\hat{F}_N^x(q_\alpha)\right)\right),\end{aligned}$$

where

$$C_{\mathbf{n}}(x)=\frac{1}{f^x(q_\alpha)}E\left(\alpha\hat{F}_D^x-\hat{F}_N^x(q_\alpha)\right).$$

Consequently, the proof of the theorem is the consequence of the following lemmas and the convergence result (7).

Lemma 4.4.6 *Under the hypotheses of Theorem 4.4.5, we have*

$$i) Var\left(\Delta_{\mathbf{i}}\right) \to \alpha(1-\alpha)$$

$$ii) \sum_{\mathbf{i},\mathbf{j}\in\mathcal{I}_{\mathbf{n}}} Cov\left(\Delta_{\mathbf{i}},\Delta_{\mathbf{j}}\right) = o\left(\widehat{\mathbf{n}}\right)$$

and

$$iii) \frac{1}{\widehat{\mathbf{n}}} var\left(\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \Delta_{\mathbf{i}}\right) \to \alpha(1-\alpha) \quad when \quad \mathbf{n}\to\infty.$$

Lemma 4.4.7 *Under the hypotheses of Theorem 4.4.5 we have*

$$\left[\frac{\widehat{\mathbf{n}}E^2K_{\mathbf{i}}}{\alpha(1-\alpha)EK_{\mathbf{i}}^2}\right]^{1/2}\left(\left[\alpha\hat{F}_D^x-\hat{F}_N^x(q_\alpha)\right]-E\left[\alpha\hat{F}_D^x-\hat{F}_N^x(q_\alpha)\right]\right)\to N(0,1).$$

Lemma 4.4.8 *Under the hypotheses H_2' and H_4', we have :*

$$E\left[\alpha\hat{F}_D^x-\hat{F}_N^x\left(q_\alpha\right)\right]=\alpha-\frac{1}{EK_{\mathbf{i}}}E\left[K\left(\frac{d,x-X)}{a_{\mathbf{n}}}\right)F^X\left(q_\alpha\right)\right]+O\left(b_{\mathbf{n}}^{b_2}\right)$$

and

$$E\left[\alpha\hat{F}_D^x-\hat{F}_N^x\left(q_\alpha\right)\right]=O\left(a_{\mathbf{n}}^{b_1}+b_{\mathbf{n}}^{b_2}\right).$$

It easy to see that, if one imposes some additional assumptions on the function $\phi_x(\cdot)$ and the bandwidth parameters ($a_{\mathbf{n}}$ and $b_{\mathbf{n}}$) we can improved our asymptotic normality by explicit asymptotic expressions of dispersion terms or by removing the bias term $C_n(x)$.

Corollary 2 *Under the hypotheses of Theorem 4.4.5 and if the bandwidth parameters ($a_{\mathbf{n}}$ and $b_{\mathbf{n}}$) and if the function $\phi_x(a_{\mathbf{n}})$ satisfies :*

$$\lim_{n\to\infty}({a_{\mathbf{n}}}^{b_1}+{b_{\mathbf{n}}}^{b_2})\sqrt{\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}})}=0 \;and\quad \lim_{n\to\infty}\frac{\phi_x(ta_{\mathbf{n}})}{\phi_x(a_{\mathbf{n}})}=\beta(t),\ \forall t\in[0,1],$$

we have

$$\left(\frac{(f^x(t_\alpha(x)))^2\delta_1^2}{\delta_2(\alpha(1-\alpha))}\right)^{(1/2)}\sqrt{n\phi_x(a_{\mathbf{n}})}\left(\widehat{t_\alpha}(x)-t_\alpha(x)\right)\to_{n\to+\infty}N(0,1),$$

where $\delta_j=-\int_0^1(K^j)'(s)\beta(s)ds,$ *for,* $j=1,\,2.$

Remark 3 If we assume that (5) is satisfied instead of (4) then it is simple to have the results of Theorems 4.4.1 and 4.4.5, the only thing that changes is condition (H5) which will be replaced by some assumption that depend on $\tilde{\beta}$.

4.5 Appendix

We first state the following lemmas which are due to Carbon et al. (1997). They are needed for the convergence of our estimates. There proofs will then be omitted.

Lemma 4.5.1 *Suppose $E_1, ..., E_r$ be sets containing m sites each with $dist(E_i,\ E_j) \geq \gamma$ for all $i \neq j$ where $1 \leq i \leq r$ and $1 \leq j \leq r$. Suppose $Z_1, ..., Z_r$ is a sequence of real-valued r.v.'s measurable with respect to $\mathcal{B}(E_1), ..., \mathcal{B}(E_r)$ respectively, and Z_i takes values in $[a, b]$. Then there exists a sequence of independent r.v.'s $Z_1^*, ..., Z_r^*$ independent of $Z_1, ..., Z_r$ such that Z_i^* has the same distribution as Z_i and satisfies*

$$\sum_{i=1}^{r} E|Z_i - Z_i^*| \leq 2r(b-a)\psi((r-1)m, m)\varphi(\gamma).$$

Lemma 4.5.2
(i) Suppose that (3) holds. Denote by $\mathcal{L}_r(\mathcal{F})$ the class of $\mathcal{F}-$measurable r.v.'s X satisfying $\|X\|_r = (E|X|^r)^{1/r} < \infty$. Suppose $X \in \mathcal{L}_r(\mathcal{B}(E))$ and $Y \in \mathcal{L}_s(\mathcal{B}(E'))$. Assume also that $1 \leq r,\ s,\ t < \infty$ and $r^{-1} + s^{-1} + t^{-1} = 1$. Then

$$(9) \qquad |EXY - EXEY| \leq C\|X\|_r\|Y\|_s\{\psi(Card(E), Card(E'))\varphi(dist(E, E'))\}^{1/t}.$$

(ii) For r.v.'s bounded with probability 1, the right-hand side of (9) can be replaced by

$$C\psi(Card(E), Card(E'))\varphi(dist(E, E')).$$

Proof of Lemma 4.4.2 : We have

$$E[\alpha\hat{F}_D^x - \hat{F}_N^x(q_\alpha(x))] = \alpha - \frac{1}{EK_\mathbf{1}}\left(E\left[K_\mathbf{1} E\left[H_\mathbf{1}(q_\alpha(x)) \mid X_\mathbf{1}\right]\right]\right).$$

We shall use the integration by parts and the usual change of variables $t = \dfrac{y-z}{b_\mathbf{n}}$, to show that

$$E[\alpha\hat{F}_D^x - \hat{F}_N^x(q_\alpha(x))] = \alpha - \frac{1}{EK_\mathbf{1}}\left(EK_\mathbf{1}\int K_2^{(1)}(t)F^{X_\mathbf{1}}((q_\alpha(x)) - b_\mathbf{n}t)dt\right).$$

Hypotheses (H2) and (H4) allow to get

$$\begin{aligned} E[\alpha\hat{F}_D^x - \hat{F}_N^x(q_\alpha(x))] &\leq \frac{1}{EK_\mathbf{1}}E\left[K_\mathbf{1}\int K_2^{(1)}(t)\left|F^x((q_\alpha(x)) - F^{X_\mathbf{1}}((q_\alpha(x)) - b_\mathbf{n}t)dt\right|\right] \\ &\leq C\left(a_\mathbf{n}^{b_1} + b_\mathbf{n}^{b_2}\right). \end{aligned}$$

■

Proof of Lemma 4.4.3 :

We have

$$\left\|\alpha\hat{F}_D^x - \hat{F}_N^x(q_\alpha(x)) - E\left[\alpha\hat{F}_D^x - \hat{F}_N^x(q_\alpha(x))\right]\right\|_p = \frac{1}{n\,EK_1}\left\|\sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}}\theta_\mathbf{i}\right\|_p .$$

where

$$\theta_\mathbf{i} = K_\mathbf{i}(\alpha - H_\mathbf{i}(q_\alpha(x))) - E\left[K_\mathbf{i}(\alpha - H_\mathbf{i}(q_\alpha(x)))\right].$$

We have $EK_\mathbf{1} = O(\phi_x(h_K))$, (because of H3), so it remains to show that

$$\left\|\sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}}\theta_\mathbf{i}\right\|_p = O(\sqrt{\widehat{\mathbf{n}}\phi_x(a_\mathbf{n})}).$$

The evaluation of this quantity is based on ideas similar to that used by Gao *et al.* (2008), see also Abdi et al (2010). More preciously, we prove the case where $p = 2m$ (for all $m \in N^*$) and we use the Hölder inequality for lower values of p.

First of all, let us notice that the notations $\theta_\mathbf{i}$ and $\xi_\mathbf{i}$ deliberately introduced above are the same as those used in Lemma 2.2 of Gao *et al.* (2008) or Abdi et al. (2010a). The proof of the lemma is completely modeled on that of Lemma 2.2 of Gao *et al.*. To make easier the understanding of the effect of the boundedness of $\theta_\mathbf{i}$ on the results, we opt to run along the lines of Gao et *al*'s proof (keeping the same notations) and give the moment results in a simpler form.
To start, note that

$$E\left[\left(\sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}}\theta_\mathbf{i}\right)^{2r}\right] = \sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} E\left[\theta_\mathbf{i}^{2m}\right] + \sum_{s=1}^{2m-1}\sum_{\nu_0+\nu_1+\ldots+\nu_s=2r} V_s(\nu_0,\nu_1,\ldots,\nu_s)$$

where $\sum_{\nu_0+\nu_1+\ldots+\nu_s=2m}$ is the summation over $(\nu_0,\nu_1,\ldots,\nu_s)$ with positive integer components satisfying $\nu_0+\nu_1+\ldots+\nu_s = 2m$ and

$$V_s(\nu_0,\nu_1,\ldots,\nu_s) = \sum_{\mathbf{i}_0\neq\mathbf{i}_1\neq\ldots\neq\mathbf{i}_s} E\left[\theta_{\mathbf{i}_0}^{\nu_0}\theta_{\mathbf{i}_1}^{\nu_1}\ldots\theta_{\mathbf{i}_s}^{\nu_s}\right]$$

where the summation $\sum_{\mathbf{i}_0\neq\mathbf{i}_1\neq\ldots\neq\mathbf{i}_s}$ is over indexes $(\mathbf{i}_0,\mathbf{i}_1,\ldots,\mathbf{i}_s)$ with each index $\mathbf{i}_j$ taking value in $\mathcal{I}_\mathbf{n}$ and satisfying $\mathbf{i}_j \neq \mathbf{i}_l$ for any $j\neq l$, $0\le j,l\le s$. By stationarity and the fact that K_2 is a distribution function, we have

$$\sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} E\left(\theta_\mathbf{i}\right)^{2m} \le C\widehat{\mathbf{n}}E\left(|\theta_\mathbf{i}|\right)^{2m} \le \widehat{\mathbf{n}}E\left(K_\mathbf{i}\right)^{2m} \le C\widehat{\mathbf{n}}\phi_x(a_\mathbf{n}).$$

To control the term $V_s(\nu_0,\nu_1,\ldots,\nu_s)$, we need to prove, for any positive integers $\nu_0,\nu_1,\nu_2,\ldots\nu_s$, the following results :
i) $E\left|\theta_{\mathbf{i}_1}^{\nu_1}\theta_{\mathbf{i}_2}^{\nu_2}\ldots\theta_{\mathbf{i}_s}^{\nu_s}\right| \le C\phi_x(a_\mathbf{n})^{1+v_s}$

ii) $V_s(\nu_0,\nu_1,...,\nu_s)=O\left((\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}}))^{s+1}\right)$, for $s=1,2,...,m-1$ and $m>1$
iii) $V_s(\nu_0,\nu_1,...,\nu_s)=O\left((\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}}))^{m}\right)$, for $m\leq s\leq 2m-1$.
To show the result $i)$, remark that the boundness of K_2 and (2) yield

$$E\left|\theta_{\mathbf{i}_1}^{\nu_1}\theta_{\mathbf{i}_2}^{\nu_2}...\theta_{\mathbf{i}_s}^{\nu_s}\right|\leq C\phi_x^{1+v_s}(a_{\mathbf{n}}).$$

Proof of ii) Note that we can write

$$V_s(\nu_0,\nu_1,...,\nu_s)=\sum_{\mathbf{i}_0\neq\mathbf{i}_1\neq...\neq\mathbf{i}_s}\left[E\left(\prod_{j=0}^{s}\theta_{\mathbf{i}_j}^{\nu_j}\right)-\prod_{j=0}^{s}E\theta_{\mathbf{i}_j}^{\nu_j}\right]+\sum_{\mathbf{i}_0\neq\mathbf{i}_1\neq...\neq\mathbf{i}_s}\prod_{j=0}^{s}E\theta_{\mathbf{i}_j}^{\nu_j}=:V_{s1}+V_{s2}.$$

Clearly, we have $|V_{s2}|\leq C\sum_{\mathbf{i}_0\neq\mathbf{i}_1\neq...\neq\mathbf{i}_s}(\phi_x(a_{\mathbf{n}}))^{s+1}\leq C\left(\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}})\right)^{s+1}$.
For the term V_{s1}, notice that

$$E\left(\prod_{j=0}^{s}\theta_{\mathbf{i}_j}^{\nu_j}\right)-\prod_{j=0}^{s}E\theta_{\mathbf{i}_j}^{\nu_j}=\sum_{l=0}^{s-1}\left(\prod_{j=0}^{l-1}E\theta_{\mathbf{i}_j}^{\nu_j}\right)\left(E\left[\prod_{j=l}^{s}\theta_{\mathbf{i}_j}^{\nu_j}\right]-E\left[\theta_{\mathbf{i}_l}^{\nu_l}\right]E\left[\prod_{j=l+1}^{s}\theta_{\mathbf{i}_j}^{\nu_j}\right]\right)$$

where we define $\prod_{j=l}^{s}.=1$ if $l>s$. Then we obtain

$$|V_{s1}|\leq\sum_{l=0}^{s-1}\left(\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}})\right)^{l}\sum_{\mathbf{i}_l\neq...\neq\mathbf{i}_s}\left|E\left[\prod_{j=l}^{s}\xi_{\mathbf{i}_j}^{\nu_j}\right]-E\left[\xi_{\mathbf{i}_l}^{\nu_l}\right]E\left[\prod_{j=l+1}^{s}\xi_{\mathbf{i}_j}^{\nu_j}\right]\right|=\sum_{l=0}^{s-1}\left(\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}})\right)^{l}V_{ls1}.$$

Let P be some positive real, we have

$$\begin{aligned}V_{ls1}&=\sum_{0<dist(\{\mathbf{i}_l\},\{\mathbf{i}_{l+1},...,\mathbf{i}_s\})\leq P}\left|E\left[\prod_{j=l}^{s}\theta_{\mathbf{i}_j}^{\nu_j}\right]-E\left[\theta_{\mathbf{i}_l}^{\nu_l}\right]E\left[\prod_{j=l+1}^{s}\theta_{\mathbf{i}_j}^{\nu_j}\right]\right|\\&+\sum_{0<dist(\{\mathbf{i}_l\},\{\mathbf{i}_{l+1},...,\mathbf{i}_s\})>P}\left|E\left[\prod_{j=l}^{s}\theta_{\mathbf{i}_j}^{\nu_j}\right]-E\left[\theta_{\mathbf{i}_l}^{\nu_l}\right]E\left[\prod_{j=l+1}^{s}\theta_{\mathbf{i}_j}^{\nu_j}\right]\right|\\&:=V_{ls11}+V_{ls12}.\end{aligned}$$

Using the result $i)$ above leads to

$$\left|E\left[\prod_{j=l}^{s}\theta_{\mathbf{i}_j}^{\nu_j}\right]-E\left[\theta_{\mathbf{i}_l}^{\nu_l}\right]E\left[\prod_{j=l+1}^{s}\theta_{\mathbf{i}_j}^{\nu_j}\right]\right|\leq\left|E\left[\prod_{j=l}^{s}\theta_{\mathbf{i}_j}^{\nu_j}\right]\right|+\left|E\left[\theta_{\mathbf{i}_l}^{\nu_l}\right]E\left[\prod_{j=l+1}^{s}\theta_{\mathbf{i}_j}^{\nu_j}\right]\right|\leq C\phi_x(a_{\mathbf{n}})^{1+v_{s+1-l}}.$$

Thus, we have

$$V_{ls11}\leq\sum_{0<dist(\{\mathbf{i}_l\},\{\mathbf{i}_{l+1},...,\mathbf{i}_s\})\leq P}C\phi_x(a_{\mathbf{n}})^{1+v_{s+1-l}}\leq C\phi_x(a_{\mathbf{n}})^{1+v_{s+1-l}}\sum_{k=1}^{P}\sum_{k\leq dist(\{\mathbf{i}_l\},\{\mathbf{i}_{l+1},...,\mathbf{i}_s\})=t<k+1}1.$$

Since $dist\left(\{\mathbf{i}_l\}, \{\mathbf{i}_{l+1}, ..., \mathbf{i}_s\}\right) = t$, it follows that there exits some $\mathbf{i}_j \in \{\mathbf{i}_{l+1}, ..., \mathbf{i}_s\}$, say $\mathbf{i}_{l+1}$, such that $dist\left(\{\mathbf{i}_l\}, \{\mathbf{i}_{l+1}\}\right) = t$, and therefore

$$\sum_{k=1}^{P} \sum_{k \le dist(\{\mathbf{i}_l\},\{\mathbf{i}_{l+1},...,\mathbf{i}_s\})=t<k+1} 1 \le \sum_{k=1}^{P} \sum_{\substack{\mathbf{i}_j \in \mathcal{I}_\mathbf{n} \\ j=l+2,...,s}} \sum_{k \le dist(\{\mathbf{i}_l\},\{\mathbf{i}_{l+1}\})=t<k+1} 1 \le$$

$$\widehat{\mathbf{n}}^{(s-1-l)} \sum_{k=1}^{P} \sum_{k \le dist(\{\mathbf{i}_l\},\{\mathbf{i}_{l+1}\})=t<k+1} 1.$$

Thus

$$V_{ls11} \le C\phi_x(a_\mathbf{n})^{1+v_{s+1-l}} \widehat{\mathbf{n}}^{(s-l)} \sum_{k=1}^{P} \sum_{k \le \|\mathbf{i}\|=t<k+1} 1 \le C\phi_x(a_\mathbf{n})^{1+v_{s+1-l}} \widehat{\mathbf{n}}^{(s-l)} \sum_{t=1}^{P} t^{N-1}$$

$$\le C\phi_x(a_\mathbf{n})^{1+v_{s+1-l}} \widehat{\mathbf{n}}^{(s-l)} P^N.$$

For the term V_{ls12}, notice that since the variables $\theta_\mathbf{i}$ are bounded, we have (see Lemma 4.5.2)

$$\left| E\left[\prod_{j=l}^{s} \theta_{\mathbf{i}_j}^{\nu_j}\right] - E\left[\theta_{\mathbf{i}_l}^{\nu_l}\right] E\left[\prod_{j=l+1}^{s} \theta_{\mathbf{i}_j}^{\nu_j}\right] \right| \le C\psi(1, s-l)\varphi(t), \quad \text{where} \quad t = dist\left(\{\mathbf{i}_l\}, \{\mathbf{i}_{l+1}, ..., \mathbf{i}_s\}\right)$$

Then, under (4) or (5), we have

$$V_{ls12} \le C \sum_{k=P+1}^{\infty} \sum_{k \le dist(\{\mathbf{i}_l\},\{\mathbf{i}_{l+1}\})=t<k+1} \varphi(t) \le C \sum_{k=P+1}^{\infty} \widehat{\mathbf{n}}^{(s-l)} \sum_{k \le \|\mathbf{i}\|=t<k+1} \varphi(t) \le C\widehat{\mathbf{n}}^{(s-l)} \sum_{t=P+1}^{\infty} t^{N-1}\varphi(t).$$

Combining the upper bounds of V_{ls11} and V_{ls12}, we have

$$\begin{aligned}
|V_{s1}| &\le C \sum_{l=0}^{s-1} \left(\widehat{\mathbf{n}}\phi_x(a_\mathbf{n})\right)^l \left[\phi_x(a_\mathbf{n})^{1+v_{s+1-l}} \widehat{\mathbf{n}}^{(s-l)} P^N + \widehat{\mathbf{n}}^{(s-l)} \sum_{t=P+1}^{\infty} t^{N-1}\varphi(t)\right] \\
&\le C \left(\widehat{\mathbf{n}}\phi_x(a_\mathbf{n})\right)^{(s+1)} \sum_{l=0}^{s-1} \left(\widehat{\mathbf{n}}\phi_x(a_\mathbf{n})\right)^{l-s-1} \left[\phi_x(a_\mathbf{n})^{1+v_{s+1-l}} \widehat{\mathbf{n}}^{(s-l)} P^N + \widehat{\mathbf{n}}^{(s-l)} \sum_{t=P+1}^{\infty} t^{N-1}\varphi(t)\right] \\
&= C \left(\widehat{\mathbf{n}}\phi_x(a_\mathbf{n})\right)^{(s+1)} \sum_{l=0}^{s-1} \left[\widehat{\mathbf{n}}^{-1} P^N \phi_x(a_\mathbf{n})^{1+v_{s+1-l}} \phi_x(a_\mathbf{n})^{l-s-1} + \widehat{\mathbf{n}}^{-1} \phi_x(a_\mathbf{n})^{(l-s-1)} \sum_{t=P+1}^{\infty} t^{N-1}\varphi(t)\right]
\end{aligned}$$

$$\le C \left(\widehat{\mathbf{n}}\phi_x(a_\mathbf{n})\right)^{(s+1)} \sum_{l=0}^{s-1} \left[\widehat{\mathbf{n}}^{-1} P^N \phi_x(a_\mathbf{n})^{1+v_{s+1-l}} \phi_x(a_\mathbf{n})^{l-s-1} + \widehat{\mathbf{n}}^{-1} \phi_x(a_\mathbf{n})^{(l-s-1)} P^{-sN} \sum_{t=P+1}^{\infty} t^{sN+N-1}\varphi(t)\right].$$

Taking $P = \phi_x(a_{\mathbf{n}})^{-1/N}$, we obtain

$$|V_{s1}| \leq C\left(\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}})\right)^{(s+1)} \sum_{l=0}^{s-1}\left[\frac{1}{\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}})}\phi_x(a_{\mathbf{n}})^{1+v_{s+1-l}}\phi_x(a_{\mathbf{n}})^{l-s-1} + \frac{1}{\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}})}(\phi_x(a_{\mathbf{n}})^{l}) \sum_{t=P+1}^{\infty} t^{sN+N-1-\delta}\right]$$

$$\leq C\left(\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}})\right)^{(s+1)}$$

since $\delta > N(p+2)$.
Proof of iii) As indicated in Gao *et al.* (2008), since the arguments are similar for any $m \leq s \leq 2m-1$, the proof is given only for $s = 2m-1$ and $N = 2$ for simplicity. In this case $V_{2m-1}(\nu_0, \nu_1, ..., \nu_{2m-1})$ is denoted W. To simplify the notations, we write $\mathbf{i} = (i, j) \in \mathbb{Z}^2$ and $\mathbf{i}_k = (i_k, j_k) \in \mathbb{Z}^2$. The main difficulty is to cope with the summation

$$\sum_{\mathbf{i}_0 \neq \mathbf{i}_1 \neq ... \mathbf{i}_{2m-1}} E\left[\theta_{\mathbf{i}_0}\theta_{\mathbf{i}_1}...\theta_{\mathbf{i}_{2m-1}}\right] = \sum_{(i_0,j_0)\neq(i_1,j_1)\neq...\neq(i_{2m-1},j_{2m-1})} E\left[\theta_{i_0j_0}\theta_{i_1j_1}...\theta_{i_{2m-1}j_{2m-1}}\right].$$

To this end, a novel ordering in $\mathbb{Z}^2$ (see Gao *et al.* (2008)), makes possible to separate the indexes into two (or more) sets whose distance is greater or smaller than P (usually larger than 1) is considered. Arrange each of the index sets $\{i_0, i_1, ..., i_{2m-1}\}$ and $\{j_0, j_1, ..., j_{2m-1}\}$ in ascending orders as (retaining the same notation for the first ordered index set for simplicity) $i_0 \leq i_1 \leq ... \leq i_{2m-1}$ and $j_{l_0} \leq j_{l_1} \leq ... \leq j_{l_{2m-1}}$, where l_k is to indicate that l_k may not be equal to k. The number of such arrangements is at most $(2m)!$. Let $\Delta i_k = i_k - i_{k-1}$ and $\Delta j_k = j_{l_k} - j_{l_k-1}$ and arrange $\{\Delta i_1, ..., \Delta i_{2m-1}\}$ and $\{\Delta j_1, ..., \Delta j_{2m-1}\}$ in decreasing orders, respectively as
$\Delta i_{a_1} \geq ... \geq \Delta i_{a_{2m-1}}$and $\Delta j_{b_1} \geq ... \geq \Delta j_{b_{2m-1}}$ Let $t_1 = \Delta i_{a_m}$, $t_2 = \Delta j_{b_m}$ and $t = \max\{t_1, t_2\}$. If $t_1 \geq t_2$ then $t = t_1$, and
$0 \leq i_{a_k} - i_{a_{k-1}} \leq t_1 \leq n_1$ for $k = m+1, ..., 2m-1$,
$0 \leq j_{l_{b_k}} - j_{l_{b_{k-1}}} \leq t \leq n_2$ for $k = m, m+1, ..., 2m-1$. Therefore

$$\text{(10)} \qquad i_{a_{k-1}} \leq i_{a_k} \leq i_{a_{k-1}} + t, \quad j_{l_{b_{k-1}}} \leq j_{b_k} \leq j_{l_{b_{k-1}}} + t \quad \text{for} \quad k = m+1, ..., 2m-1.$$

We arrange $\mathbf{i}_0 \neq \mathbf{i}_1 \neq ... \neq \mathbf{i}_{2m-1}$ according to the order of $i_0 \leq i_1 \leq ... \leq i_{2m-1}$. If $t_1 < t_2$, arrange according to the order of $j_{l_0} \leq j_{l_1} \leq ... \leq j_{l_{2m-1}}$ and the proof is similar.
Let $\mathcal{I} = \{i_1, ..., i_{2m-1}\}$, $\mathcal{I}_a = \{i_{a_1}, ..., i_{a_m}\}$, $\mathcal{I}_a^c = \mathcal{I} - \mathcal{I}_a = \{i_{a_{m+1}}, ..., i_{a_{2m-1}}\}$,
$\mathcal{J} = \{j_{l_1}, ..., j_{l_{2m-1}}\}$, $\mathcal{J}_b = \{j_{l_{b_1}}, ..., j_{l_{b_m}}\}$ and $\mathcal{J}_b^c = \mathcal{J} - \mathcal{J}_b = \{j_{l_{b_{m+1}}}, ..., j_{l_{b_{2m-1}}}\}$.
Remark that $(i_{a_1}, ..., i_{a_{2m-1}})$ and $(j_{l_{b_1}}, ..., j_{l_{b_{2m-1}}})$ are permutations of respectively $\mathcal{I}$ and $\mathcal{J}$.

Then, from (10), $t = i_{a_m} - i_{a_{m-1}}$ and $t \geq j_{l_{b_m}} - j_{l_{b_{m-1}}}$, we deduce that

$$\begin{aligned}
W &= \sum_{\mathbf{i}_0 \neq \mathbf{i}_1 \neq \ldots \mathbf{i}_{2m-1}} E\left[\theta_{\mathbf{i}_0}\theta_{\mathbf{i}_1}\ldots\theta_{\mathbf{i}_{2m-1}}\right] \\
&\leq C \sum_{1 \leq i_0 \leq i_1 \leq \ldots \leq i_{2m-1} \leq n_1} \sum_{1 \leq j_{l_0} \leq j_{l_1} \leq \ldots \leq j_{l_{2m-1}} \leq n_2} \left|E\left[\theta_{i_0 j_0}\theta_{i_1 j_1}\ldots\theta_{i_{2m-1} j_{2m-1}}\right]\right| \\
&\leq C \sum_{t=1}^{\max(n_1,n_2)} \sum_{i_0=1}^{n_1} \sum_{\substack{i=1 \\ i \in \mathcal{I}_a - \{i_{a_m}\}}}^{n_1} \sum_{\substack{i_{a_k} = i_{a_{k-1}} \\ k=m+1,\ldots,2m-1}}^{i_{a_{k-1}}+t} \sum_{j_{l_0}=1}^{n_2} \sum_{\substack{j=1 \\ j \in \mathcal{J}_b - \{j_{l_{b_m}}\}}}^{n_2} \sum_{\substack{j_{l_{b_k}} = j_{l_{b_{k-1}}} \\ k=m,m+1,\ldots,2m-1}}^{j_{l_{b_{k-1}}}+t} \left|E\left[\theta_{\mathbf{i}_0}\theta_{\mathbf{i}_1}\ldots\theta_{\mathbf{i}_{2m-1}}\right]\right|.
\end{aligned}$$

Take a positive constant P such that $1 \leq P \leq \max(n_1, n_2)$ and divide the right hand side of the previous inequality into two parts denoted by W_1 and W_2 according to $1 \leq t \leq P$ and $t > P$. Then $W \leq W_1 + W_2$. In one hand, use the result $i)$ with $s = 2m$, and get

$$\begin{aligned}
W_1 &= C \sum_{t=1}^{P} \sum_{i_0=1}^{n_1} \sum_{\substack{i=1 \\ i \in \mathcal{I}_a - \{i_{a_m}\}}}^{n_1} \sum_{\substack{i_{a_k} = i_{a_{k-1}} \\ k=m+1,\ldots,2m-1}}^{i_{a_{k-1}}+t} \sum_{j_{l_0}=1}^{n_2} \sum_{\substack{j=1 \\ j \in \mathcal{J}_b - \{j_{l_{b_m}}\}}}^{n_2} \sum_{\substack{j_{l_{b_k}} = j_{l_{b_{k-1}}} \\ k=m,m+1,\ldots,2m-1}}^{j_{l_{b_{k-1}}}+t} \left|E\left[\theta_{\mathbf{i}_0}\theta_{\mathbf{i}_1}\ldots\theta_{\mathbf{i}_{2m-1}}\right]\right| \\
&\leq C \sum_{t=1}^{P} \sum_{i_0=1}^{n_1} \sum_{\substack{i=1 \\ i \in \mathcal{I}_a - \{i_{a_m}\}}}^{n_1} \sum_{\substack{i_{a_k} = i_{a_{k-1}} \\ k=m+1,\ldots,2m-1}}^{i_{a_{k-1}}+t} \sum_{j_{l_0}=1}^{n_2} \sum_{\substack{j=1 \\ j \in \mathcal{J}_b - \{j_{l_{b_m}}\}}}^{n_2} \sum_{\substack{j_{l_{b_k}} = j_{l_{b_{k-1}}} \\ k=m,m+1,\ldots,2m-1}}^{j_{l_{b_{k-1}}}+t} \phi_x(a_{\mathbf{n}})^{1+v_{2m}} \\
&\leq C(n_1 n_2)^m \sum_{t=1}^{P} t^{2m-1} \phi_x(a_{\mathbf{n}})^{1+v_{2m}} \leq C(n_1 n_2)^m P^{2m} \phi_x(a_{\mathbf{n}})^{1+v_{2m}}.
\end{aligned}$$

In another hand, assume that neither i_1 nor i_{2m-1} belongs to $\mathcal{I}_a$ (if i_1 or i_{2m-1} is in $\mathcal{I}_a$ the proof is similar). In this case, $\mathcal{I}_a$ is a subset of size m chosen from the $2m-3$ remaining indexes (besides i_0, i_1 and i_{2m-1}). As raised by Gao et al., this is due to the fact that there must be two successive indexes because there are not enough elements in the set of remaining indices to allow a gap between every two elements of $\mathcal{I}_a$. The first components of $\mathbf{i}_j$'s are ordered as $i_0 \leq i_1 \leq \ldots \leq i_{k^*-1} \leq i_{k^*} \leq i_{k^*+1} \leq \ldots \leq i_{2m-1}$ for some $k^* \geq 1$ and $\Delta i_j = i_j - i_{j-1}$. Then we have, either
$dist(\{\mathbf{i}_0, \ldots, \mathbf{i}_{k^*-1}\}, \{\mathbf{i}_{k^*}\}) \geq \Delta i_{k^*} \geq t$, $dist(\{\mathbf{i}_{k^*}\}, \{\mathbf{i}_{k^*+1}, \ldots, \mathbf{i}_{2m-1}\}) \geq \Delta i_{k^*+1} \geq t$ and $dist(\{\mathbf{i}_0, \ldots, \mathbf{i}_{k^*-1}\}, \{\mathbf{i}_{k^*}, \ldots, \mathbf{i}_{2m-1}\}) \geq \Delta i_{k^*+1} \geq t$, or $dist(\{\mathbf{i}_0\}, \{\mathbf{i}_1, \ldots, \mathbf{i}_{2m-1}\}) \geq \Delta i_1 \geq t$ or $dist(\{\mathbf{i}_0, \ldots, \mathbf{i}_{2m-2}\}, \{\mathbf{i}_{2m-1}\}) \geq \Delta i_{2m-1} \geq t$. Let $A_{\mathbf{i}_{k^*-1}} = \theta_{\mathbf{i}_0}\theta_{\mathbf{i}_1}\ldots\theta_{\mathbf{i}_{k^*-1}}$ and $B_{\mathbf{i}_{k^*+1}} =$

$\theta_{\mathbf{i}_{k^*+1}}...\theta_{\mathbf{i}_{2m-1}}$, then for the case of i_{k^*} and i_{k^*+1} in $\mathcal{I}_a$, we have

$$\begin{aligned}
\left|E\left[\theta_{\mathbf{i}_0}\theta_{\mathbf{i}_1}...\theta_{\mathbf{i}_{2m-1}}\right]\right| &= \left|E\left[A_{\mathbf{i}_{k^*-1}}\theta_{\mathbf{i}_{k^*}}B_{\mathbf{i}_{k^*+1}}\right]\right| \\
&\leq \left|E\left[\left(A_{\mathbf{i}_{k^*-1}} - EA_{\mathbf{i}_{k^*-1}}\right)\left(\theta_{\mathbf{i}_{k^*}}B_{\mathbf{i}_{k^*+1}} - E\theta_{\mathbf{i}_{k^*}}B_{\mathbf{i}_{k^*+1}}\right)\right]\right| \\
&+ \left|E\left[A_{\mathbf{i}_{k^*-1}}\right]E\left[\left(\theta_{\mathbf{i}_{k^*}}B_{\mathbf{i}_{k^*+1}}\right)\right]\right| \\
&= \left|Cov\left(A_{\mathbf{i}_{k^*-1}}, \theta_{\mathbf{i}_{k^*}}B_{\mathbf{i}_{k^*+1}}\right)\right| + \left|E\left[A_{\mathbf{i}_{k^*-1}}\right]\right|\left|Cov\left(\theta_{\mathbf{i}_{k^*}}, B_{\mathbf{i}_{k^*+1}}\right)\right| \\
&\leq C\varphi(t) + C\phi_x(a_{\mathbf{n}})^{1+v_{k^*}}\varphi(t) \leq C\varphi(t).
\end{aligned}$$

Thus

$$\begin{aligned}
W_2 &= C\sum_{t=P+1}^{\max(n_1,n_2)}\sum_{i_0=1}^{n_1}\sum_{\substack{i=1\\ i\in\mathcal{I}_a-\{i_{a_m}\}}}^{n_1}\sum_{\substack{i_{a_k}=i_{a_{k-1}}\\ k=m+1,...,2m-1}}^{i_{a_{k-1}}+t}\sum_{j_{l_0}=1}^{n_2}\sum_{\substack{j=1\\ j\in\mathcal{J}_b-\{j_{l_{b_m}}\}}}^{n_2}\sum_{\substack{j_{l_{b_k}}=j_{l_{b_{k-1}}}\\ k=m,m+1,...,2m-1}}^{j_{l_{b_{k-1}}}+t}\left|E\left[\theta_{\mathbf{i}_0}\theta_{\mathbf{i}_1}...\theta_{\mathbf{i}_{2m-1}}\right]\right| \\
&\leq C(n_1n_2)^m\sum_{t=P+1}^{\infty}t^{2m-1}\varphi(t).
\end{aligned}$$

It follows that

$$\begin{aligned}
W \leq W_1 + W_2 &\leq C(n_1n_2)^m P^{2m}\phi_x(a_{\mathbf{n}})^{1+v_{2m}} + C(n_1n_2)^m\sum_{t=P+1}^{\infty}t^{2m-1}\varphi(t) \\
&\leq (n_1n_2)^m\left(P^{2m}\phi_x(a_{\mathbf{n}})^{(1+v_{2m})} + P^{2m-1-\delta}\right).
\end{aligned}$$

For general N, we obtain by similar arguments

$$W \leq C(\widehat{\mathbf{n}})^m\left(P^{Nm}\phi_x(a_{\mathbf{n}})^{(1+v_s)} + P^{Nm-1-\delta}\right).$$

Taking $P = \phi_x(a_{\mathbf{n}})^{-(1+v_{Nm})/1+\delta)}$, we get

$$W \leq C(\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}}))^m\left(\widehat{\mathbf{n}}^{-1}\phi_x(a_{\mathbf{n}})^{-Nm-1+(1+v_{Nm})} + (\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}}))^{-1}\sum_{t=P+1}^{\infty}t^{Nm-1-\delta}\right) \leq C(\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}}))^m$$

because $\delta > N(p+2)$. This ends the proof of the lemma.

Proof of Lemma 4.4.4

We have for all $\epsilon < 1$,

$$\begin{aligned}
P\left(\widehat{F}_D^x = 0\right) &\leq P\left(\widehat{F}_D^x \leq 1-\epsilon\right) \\
&\leq P\left(|\widehat{F}_D^x - E[\widehat{F}_D^x]| \geq \epsilon\right).
\end{aligned}$$

Markov's inequality allows to get, for any $p > 0$,

$$P\left(|\widehat{F}_D^x - E[\widehat{F}_D^x]| \geq \epsilon\right) \leq \frac{E\left[|\widehat{F}_D^x - E[\widehat{F}_D^x]|^p\right]}{\epsilon^p}.$$

So
$$\left(P\left(\widehat{F}_D^x=0\right)\right)^{1/p}=O\left(\left\|\widehat{F}_D^x-E[\widehat{F}_D^x]\right\|_p\right).$$
The computation of $\left\|\widehat{F}_D^x-E[\widehat{F}_D^x]\right\|_p$ can be done by following same arguments as those used to prove Lemma 4.4.3. This yields the proof. ■

Proof of Lemma 4.4.6

Let us calculate the variance $Var\left(\Delta_{\mathbf{i}}\right)$. We have :

$$\begin{aligned}
Var\left(\Delta_{\mathbf{i}}\right) &= \frac{1}{EK_{\mathbf{i}}^2}\left[EK_{\mathbf{i}}^2\left(\alpha-H_{\mathbf{i}}\left(q_\alpha\right)\right)^2-\left(EK_{\mathbf{i}}\left(\alpha-H_{\mathbf{i}}\left(q_\alpha\right)\right)\right)^2\right]\\
&= \frac{1}{EK_{\mathbf{i}}^2}EK_{\mathbf{i}}^2\left(\alpha-H_{\mathbf{i}}\left(q_\alpha\right)\right)^2-\frac{\left(EK_{\mathbf{i}}\right)^2}{EK_{\mathbf{i}}^2}\left[E\frac{K_{\mathbf{i}}\left(H_{\mathbf{i}}\left(q_\alpha\right)-\alpha\right)}{EK_{\mathbf{i}}}\right]^2\\
&= A_1-A_2
\end{aligned}$$

Let us first consider A_2. We deduce from the hypothesis H_3 that there exist two positive constants C and C' such that : $C\phi_x(a_{\mathbf{n}})\le EK_{\mathbf{i}}^r\le C'\phi_x(a_{\mathbf{n}})$, $r>1$, thus $\frac{(EK_{\mathbf{i}})^2}{EK_{\mathbf{i}}^2}=o(1)$. If we take the conditional expectation with respect to X, we get :

$$\left|E\left[\frac{K_{\mathbf{i}}\left(H_{\mathbf{i}}\left(q_\alpha\right)-\alpha\right)}{EK_{\mathbf{i}}}\right]\right|=\left|E\frac{K_{\mathbf{i}}}{EK_{\mathbf{i}}}\left[E\left(H_{\mathbf{i}}\left(q_\alpha\right)|X\right)-\alpha\right]\right|\le E\frac{K_{\mathbf{i}}}{EK_{\mathbf{i}}}\left|E\left(H_{\mathbf{i}}\left(q_\alpha\right)|X_{\mathbf{i}}\right)-\alpha\right|.$$

It is easy to see that by hypothesis $H_4'(ii)$:

$$\left|E\left(H_{\mathbf{i}}\left(q_\alpha\right)|X\right)-\alpha\right|=\left|E\left(H_{\mathbf{i}}\left(q_\alpha\right)|X\right)-F^x(q_\alpha)\right|\le C\left(a_{\mathbf{n}}^{b_1}+b_{\mathbf{n}}^{b_2}\int_{\mathbb{R}}|t|^{b_2}\left|K_2^{(1)}(t)\right|dt\right),$$

$$\left|E\frac{K_{\mathbf{i}}\left(H_{\mathbf{i}}\left(q_\alpha\right)-\alpha\right)}{EK_{\mathbf{i}}}\right|=O\left(a_{\mathbf{n}}^{b_1}+b_{\mathbf{n}}^{b_2}\right).$$

Then we deduce that A_2 tends to 0. Concerning A_1, we have :

$$\left(\alpha-H_{\mathbf{i}}\left(q_\alpha\right)\right)^2=\left(H_{\mathbf{i}}^2\left(q_\alpha\right)-\alpha\right)-2\alpha\left(H_{\mathbf{i}}\left(q_\alpha\right)-\alpha\right)+\alpha-\alpha^2.$$

Then, we can write :

$$A_1=\frac{1}{EK_{\mathbf{i}}^2}\left[EK_{\mathbf{i}}^2\left(H_{\mathbf{i}}^2\left(q_\alpha\right)-\alpha\right)-2\alpha EK_{\mathbf{i}}^2\left(H_{\mathbf{i}}\left(q_\alpha\right)-\alpha\right)\right]+\alpha\left(1-\alpha\right).$$

The conditional expectation with respect to $X_{\mathbf{i}}$, permits to obtain :

$$A_1=E\frac{K_{\mathbf{i}}^2}{EK_{\mathbf{i}}^2}\left[E\left(H_{\mathbf{i}}^2\left(q_\alpha\right)|X_{\mathbf{i}}\right)-\alpha\right]-2\alpha E\frac{K_{\mathbf{i}}^2}{EK_{\mathbf{i}}^2}\left[E\left(H_{\mathbf{i}}\left(q_\alpha\right)|X_{\mathbf{i}}\right)-\alpha\right]+\alpha\left(1-\alpha\right).$$

A same argument as above, gives

$$\left|E\frac{K_{\mathbf{i}}^{2}}{EK_{\mathbf{i}}^{2}}\left[E\left(H_{\mathbf{i}}\left(q_{\alpha}\right)|X_{\mathbf{i}}\right)-\alpha\right]\right|=O\left(a_{\mathbf{n}}^{b_1}+b_{\mathbf{n}}^{b_2}\right).$$

It remains to show that

$$\left|E\left(H_{\mathbf{i}}^{2}\left(q_{\alpha}\right)|X_{\mathbf{i}}\right)-\alpha\right|=O\left(a_{\mathbf{n}}^{b_1}+b_{\mathbf{n}}^{b_2}\right).$$

By an integration by part and hypotheses H_2' and H_4', we have :

$$\begin{aligned}
\left|E\left(H_{\mathbf{i}}^{2}\left(q_{\alpha}\right)|X_{\mathbf{i}}\right)-\alpha\right| &= \left|\int_{\mathbb{R}}K_2^2\left(\frac{q_\alpha-z}{b_{\mathbf{n}}}\right)f^{X_{\mathbf{i}}}(z)dz-F^x(q_\alpha)\right|\\
&= \left|\int_{\mathbb{R}}2K_2(t)K_2^{(1)}(t)\left(F^{X_{\mathbf{i}}}(q_\alpha-b_{\mathbf{n}}t)-F^x(q_\alpha)\right)dt\right|\\
&\leq a_{\mathbf{n}}^{b_1}\int_{\mathbb{R}}2K_2(t)K_2^{(1)}(t)dt+b_{\mathbf{n}}^{b_2}\int_{\mathbb{R}}2K_2(t)|t|^{b_2}K_2^{(1)}(t)dt\\
&\leq Ca_{\mathbf{n}}^{b_1}+b_{\mathbf{n}}^{b_2}\int_{\mathbb{R}}2|t|^{b_2}K_2^{(1)}(t)dt=O\left(a_{\mathbf{n}}^{b_1}+b_{\mathbf{n}}^{b_2}\right).
\end{aligned}$$

We deduce from above that A_1 converges to $\alpha\left(1-\alpha\right)$; then

$$Var\left(\Delta_{\mathbf{i}}\right)\rightarrow\alpha\left(1-\alpha\right).$$

Let us focus now on the covariance term. We consider

$$E_1=\{\mathbf{i},\mathbf{j}\in\mathcal{I}_{\mathbf{n}}:0<\|\mathbf{i}-\mathbf{j}\|\leq c_{\mathbf{n}}\},$$

$$E_2=\{\mathbf{i},\mathbf{j}\in\mathcal{I}_{\mathbf{n}}:\|\mathbf{i}-\mathbf{j}\|>c_{\mathbf{n}}\}.$$

We have

$$\begin{aligned}
Cov\left(\Delta_{\mathbf{i}},\Delta_{\mathbf{j}}\right) &= E\Delta_{\mathbf{i}}\Delta_{\mathbf{j}}\\
&= \frac{1}{EK_{\mathbf{i}}^2}\left[EK_{\mathbf{i}}K_{\mathbf{j}}\left(\alpha-H_{\mathbf{i}}(q_\alpha)\right)\left(\alpha-H_{\mathbf{j}}(q_\alpha)\right)-\left(EK_{\mathbf{i}}\left(\alpha-H_{\mathbf{i}}(q_\alpha)\right)\right)^2\right]\\
&\leq \frac{1}{EK_{\mathbf{i}}^2}\left[EK_{\mathbf{i}}K_{\mathbf{j}}\left(\alpha-H_{\mathbf{i}}(q_\alpha)\right)\left(\alpha-H_{\mathbf{j}}(q_\alpha)\right)\right]+\frac{1}{EK_{\mathbf{i}}^2}\left[EK_{\mathbf{i}}\left(\alpha-H_{\mathbf{i}}(q_\alpha)\right)\right]^2.
\end{aligned}$$

The conditional expectation with respect to $X_{\mathbf{i}}$, gives :

$$\frac{1}{EK_{\mathbf{i}}^2}\left[EK_{\mathbf{i}}\left(\alpha-H_{\mathbf{i}}(q_\alpha)\right)\right]^2=\frac{1}{EK_{\mathbf{i}}^2}\left|EK_{\mathbf{i}}\left(\alpha-E(H_{\mathbf{i}}(q_\alpha)|X_{\mathbf{i}})\right)\right|^2\leq\frac{1}{EK_{\mathbf{i}}^2}\left[EK_{\mathbf{i}}\left|E\left(H_{\mathbf{i}}(q_\alpha)|X_{\mathbf{i}}\right)-\alpha\right|\right]^2.$$

Recall that

$$\left|E\left(H_{\mathbf{i}}(q_\alpha)|X\right)-\alpha\right|=O\left(a_{\mathbf{n}}^{b_1}+b_{\mathbf{n}}^{b_2}\right);\text{then}\left|E\left(H_{\mathbf{i}}(q_\alpha)|X\right)-\alpha\right|\leq C.$$

So
$$\frac{1}{EK_{\mathbf{i}}^2}\left[EK_{\mathbf{i}}\left(\alpha - H_{\mathbf{i}}(q_\alpha)\right)\right]^2 \le C\phi_x(a_{\mathbf{n}}).$$
Since K_1 is bounded, we get
$$\frac{1}{EK_{\mathbf{i}}^2}\left[EK_{\mathbf{i}}K_{\mathbf{j}}\left(\alpha - H_{\mathbf{i}}(q_\alpha)\right)\left(\alpha - H_{\mathbf{j}}(q_\alpha)\right)\right] \le C\frac{1}{EK_{\mathbf{i}}^2}\left[EK_{\mathbf{i}}K_{\mathbf{j}}\right] \le$$
$$C\frac{1}{EK_{\mathbf{i}}^2}P\left[(X_{\mathbf{i}}, X_{\mathbf{j}}) \in B(x, a_{\mathbf{n}}) \times B(x, a_{\mathbf{n}})\right].$$

Then we deduce from (2)
$$\begin{aligned} & \frac{1}{EK_{\mathbf{i}}^2}\left[EK_{\mathbf{i}}K_{\mathbf{j}}\left(\alpha - H_{\mathbf{i}}(q_\alpha)\right)\left(\alpha - H_{\mathbf{j}}(q_\alpha)\right)\right] \\ \le\ & C\frac{1}{EK_{\mathbf{i}}^2}(\phi_x(a_{\mathbf{n}}))^{1+v_2} \le (\phi_x(a_{\mathbf{n}}))^{v_2}. \end{aligned}$$
Then, we have since $v_2 > 1$: $Cov\left(\Delta_{\mathbf{i}}, \Delta_{\mathbf{j}}\right) \le C(\phi_x(a_{\mathbf{n}}) + (\phi_x(a_{\mathbf{n}}))^{v_2}) \le C(\phi_x(a_{\mathbf{n}}))$ and $\sum_{E_1} Cov\left(\Delta_{\mathbf{i}}, \Delta_{\mathbf{j}}\right) \le C\widehat{\mathbf{n}}c_{\mathbf{n}}^N\phi_x(a_{\mathbf{n}})$.
Lemma 4.5.2 and $|\Delta_{\mathbf{i}}| \le C\phi_x(a_{\mathbf{n}})^{-1/2}$, permit to write that :
$$\left|Cov\left(\Delta_{\mathbf{i}}, \Delta_{\mathbf{j}}\right)\right| \le C\phi_x(a_{\mathbf{n}})^{-1}\varphi\left(\|\mathbf{i} - \mathbf{j}\|\right)$$
and
$$\begin{aligned} \sum_{E_2} Cov\left(\Delta_{\mathbf{i}}, \Delta_{\mathbf{j}}\right) &\le C\phi_x(a_{\mathbf{n}})^{-1}\sum_{(\mathbf{i},\mathbf{j})\in E_2}\varphi\left(\|\mathbf{i} - \mathbf{j}\|\right) \\ &\le C\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}})^{-1}\sum_{\mathbf{i}:\|\mathbf{i}\|>c_{\mathbf{n}}}\varphi\left(\|\mathbf{i}\|\right) \\ &\le C\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}})^{-1}c_{\mathbf{n}}^{-\delta}\sum_{\mathbf{i}:\|\mathbf{i}\|>c_{\mathbf{n}}}\|\mathbf{i}\|^{\delta}\varphi\left(\|\mathbf{i}\|\right). \end{aligned}$$

Finally, for $\delta > 0$ we have :
$$\sum Cov\left(\Delta_{\mathbf{i}}, \Delta_{\mathbf{j}}\right) \le \left(C\widehat{\mathbf{n}}c_{\mathbf{n}}^N\phi_x(a_{\mathbf{n}}) + C\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}})^{-1}c_{\mathbf{n}}^{-\delta}\sum_{\mathbf{i}:\|\mathbf{i}\|>c_{\mathbf{n}}}\|\mathbf{i}\|^{\delta}\varphi\left(\|\mathbf{i}\|\right)\right).$$
Let $c_{\mathbf{n}} = \phi_x(a_{\mathbf{n}})^{-1/N}$, then we have :
$$\sum Cov\left(\Delta_{\mathbf{i}}, \Delta_{\mathbf{j}}\right) \le \left(C\widehat{\mathbf{n}} + C\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}})^{\delta/N-1}\sum_{\mathbf{i}:\|\mathbf{i}\|>c_{\mathbf{n}}}\|\mathbf{i}\|^{\delta}\varphi\left(\|\mathbf{i}\|\right)\right).$$

Hence, we obtain that

$$\sum Cov\left(\Delta_{\mathbf{i}}, \Delta_{\mathbf{j}}\right) = o\left(\widehat{\mathbf{n}}\right).$$

In conclusion, we have

$$\frac{1}{\widehat{\mathbf{n}}} var\left(\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \Delta_{\mathbf{i}}\right) = \left(var\left(\Delta_{\mathbf{i}}\right) + \frac{1}{\widehat{\mathbf{n}}}\sum_{\mathbf{i},\mathbf{j}\in\mathcal{I}_{\mathbf{n}}} Cov\left(\Delta_{\mathbf{i}}, \Delta_{\mathbf{j}}\right)\right) \to \alpha(1-\alpha) \text{ when } \mathbf{n}\to\infty.$$

This yields the proof.

Proof of Lemma 4.4.7

Let

$$S_{\mathbf{n}} = \sum_{\substack{j_k=1\\ k=1,\ldots,N}}^{n_k} \Delta_{\mathbf{j}} \text{ with } \Delta_{\mathbf{j}} = \frac{1}{\sqrt{EK_{\mathbf{i}}}}\left[\alpha K_{\mathbf{i}} - K_{\mathbf{i}} H_{\mathbf{i}}\left(q_\alpha\right) - E\left(\alpha K_{\mathbf{i}} - K_{\mathbf{i}} H_{\mathbf{i}}\left(q_\alpha\right)\right)\right].$$

Then, we can write

$$\left[\frac{\widehat{\mathbf{n}} E^2 K_{\mathbf{i}}}{\alpha(1-\alpha) E K_{\mathbf{i}}^2}\right]^{1/2}\left(\left[\alpha \hat{F}_D^x - \hat{F}_N^x(q_\alpha)\right] - E\left[\alpha \hat{F}_D^x - \hat{F}_N^x(q_\alpha)\right]\right) = \left(\widehat{\mathbf{n}}\alpha(1-\alpha)\right)^{-1/2} S_{\mathbf{n}}.$$

Consider the same spatial block decomposition (due to Tran (1990)) as Lemma 4.4.3, with $q_{\mathbf{n}} = o\left(\left[\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}})^{(1+2N)}\right]^{1/(2N)}\right)$, $m_{\mathbf{n}} = \left[(\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}}))^{1/(2N)}/s_{\mathbf{n}}\right]$ where $s_{\mathbf{n}} = o\left(\left[\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}})^{1+2N}\right]^{1/(2N)} q_{\mathbf{n}}^{-1}\right)$. Then, we have

$$S_{\mathbf{n}} = \sum_{i=1}^{2^N} T(\mathbf{n}, x, i),$$

where

$$T(\mathbf{n}, x, i) = \sum_{\mathbf{j}\in\mathcal{J}} U(i, \mathbf{n}, x, \mathbf{j}).$$

Hence,

$$S_{\mathbf{n}} / \left(\widehat{\mathbf{n}}\alpha(1-\alpha)\right)^{1/2} = T\left(\mathbf{n}, x, 1\right) / \left(\widehat{\mathbf{n}}\alpha(1-\alpha)\right)^{1/2} + \sum_{i=2}^{2N} T\left(\mathbf{n}, x, i\right) / \left(\widehat{\mathbf{n}}\alpha(1-\alpha)\right)^{1/2}.$$

Thus the proof of the asymptotic normality of $\left(\widehat{\mathbf{n}}\alpha(1-\alpha)\right)^{-1/2} S_{\mathbf{n}}$ is reduced to the proofs of the following results

$$Q_1 \equiv \left| E\exp\left[iuT(\mathbf{n}, x, 1)\right] - \prod_{\substack{j_k=0\\ k=1,\ldots,N}}^{r_k-1} E\exp\left[iuU(1, \mathbf{n}, x, \mathbf{j})\right]\right| \to 0 \tag{11}$$

(12) $$Q_2 \equiv \widehat{\mathbf{n}}^{-1} E\Big(\sum_{i=2}^{2^N} T(\mathbf{n}, x, i)\Big)^2 \to 0$$

(13) $$Q_3 \equiv \widehat{\mathbf{n}}^{-1} \sum_{\mathbf{j}\in\mathcal{J}} E\Big[U(1, \mathbf{n}, x, \mathbf{j})\Big]^2 \to \alpha(1-\alpha)$$

(14) $$Q_4 \equiv \widehat{\mathbf{n}}^{-1} \sum_{\mathbf{j}\in\mathcal{J}} E\Big[(U(1, \mathbf{n}, x, \mathbf{j}))^2 \mathbf{1}_{\{|U(1,\mathbf{n},x,\mathbf{j})|>\epsilon(\alpha(1-\alpha)\widehat{\mathbf{n}})^{1/2}\}}\Big] \to 0 \text{ for all } \epsilon > 0.$$

Proof of (11) : Let us numerate the $M = \prod_{k=1}^N r_k = \widehat{\mathbf{n}}(m_{\mathbf{n}} + q_{\mathbf{n}})^{-N} \leq \widehat{\mathbf{n}} m_{\mathbf{n}}^{-N}$ random variables $U(1, \mathbf{n}, x, \mathbf{j}); \mathbf{j} \in \mathcal{J}$ in the arbitrary way $\tilde{U_1}, ..., \tilde{U_M}$. For $\mathbf{j} \in \mathcal{J}$, let

$$I(1, \mathbf{n}, x, \mathbf{j}) = \{\mathbf{i} : j_k(m_{\mathbf{n}} + q_{\mathbf{n}}) + 1 \leq i_k \leq j_k(m_{\mathbf{n}} + q_{\mathbf{n}}) + m_{\mathbf{n}} \quad ; k = 1, ..N\}$$

then we have $U(1, \mathbf{n}, x, \mathbf{j}) = \sum_{\mathbf{i}\in I(1,\mathbf{n},x,\mathbf{j})} \Delta_{\mathbf{i}}$. Note that each of the sets of site $I(1, \mathbf{n}, x, \mathbf{j})$ contains $m_{\mathbf{n}}^N$, these sets are distant of $m_{\mathbf{n}}$ at least. Let us apply the lemma of Volkonski and Rozanov (1959) to the variable $\Big(\exp(iu\tilde{U}_1), ..., \exp(iu\tilde{U}_M)\Big)$. Thee fact that $\Big|\prod_{s=j+1}^{M} \exp[iu\tilde{U}_s]\Big| \leq 1$, implies :

$$\begin{aligned}
Q_1 &= \Big|E \exp\big[iuT(\mathbf{n}, x, 1)\big] - \prod_{\substack{j_k=0\\k=1,\dots,N}}^{r_k-1} E \exp\big[iuU(1, \mathbf{n}, x, \mathbf{j})\big]\Big| \\
&= \left|E \prod_{\substack{j_k=0\\k=1,\dots,N}}^{r_k-1} \exp\big[iuU(1, \mathbf{n}, x, \mathbf{j})\big] - \prod_{\substack{j_k=0\\k=1,\dots,N}}^{r_k-1} E \exp\big[iuU(1, \mathbf{n}, x, \mathbf{j})\big]\right| \\
&\leq \sum_{k=1}^{M-1} \sum_{j=k+1}^{M} \Big|E\Big(\exp[iu\tilde{U}_k] - 1\Big)\Big(\exp[iu\tilde{U}_j] - 1\Big) \prod_{s=j+1}^{M} \exp[iu\tilde{U}_s] \\
&\qquad\qquad - E\Big(\exp[iu\tilde{U}_k] - 1\Big) E\Big(\exp[iu\tilde{U}_j] - 1\Big) \prod_{s=j+1}^{M} \exp[iu\tilde{U}_s]\Big| \\
&= \sum_{k=1}^{M-1} \sum_{j=k+1}^{M} \Big|E\Big(\exp[iu\tilde{U}_k] - 1\Big)\Big(\exp[iu\tilde{U}_j] - 1\Big) - E\Big(\exp[iu\tilde{U}_k] - 1\Big) E\Big(\exp[iu\tilde{U}_j] - 1\Big)\Big| \\
&\qquad\qquad \times \Big|\prod_{s=j+1}^{M} \exp[iu\tilde{U}_s]\Big| \\
&\leq \sum_{k=1}^{M-1} \sum_{j=k+1}^{M} \Big|E\Big(\exp[iu\tilde{U}_k] - 1\Big)\Big(\exp[iu\tilde{U}_j] - 1\Big) - E\Big(\exp[iu\tilde{U}_k] - 1\Big) E\Big(\exp[iu\tilde{U}_j] - 1\Big)\Big|.
\end{aligned}$$

Let $\tilde{I}_j$ be the set of sites among the $I(1,\mathbf{n},x,\mathbf{j})$ such that $\tilde{U}_j = \sum_{\mathbf{i}\in\tilde{I}(\mathbf{j})} \Delta_{\mathbf{i}}$. The lemma of Carbon et al. (1997) and assumption (3), give :

$$\left| E\left(\exp[iu\tilde{U}_k]-1\right)\left(\exp[iu\tilde{U}_j]-1\right) - E\left(\exp[iu\tilde{U}_k]-1\right)E\left(\exp[iu\tilde{U}_j]-1\right)\right| \le C\varphi\left(d(\tilde{I}_j,\tilde{I}_k)\right)m_{\mathbf{n}}^N.$$

Then

$$\begin{aligned} Q_1 &\le Cm_{\mathbf{n}}^N \sum_{k=1}^{M-1}\sum_{j=k+1}^{M} \varphi\left(d(\tilde{I}_j,\tilde{I}_k)\right) \\ &\le Cm_{\mathbf{n}}^N M \sum_{k=2}^{M} \varphi\left(d(\tilde{I}_1,\tilde{I}_k)\right) \\ &\le Cm_{\mathbf{n}}^N M \sum_{i=1}^{\infty} \sum_{k: iq_{\mathbf{n}}\le d(\tilde{I}_1,\tilde{I}_k)<(i+1)q_{\mathbf{n}}} \varphi\left(d(\tilde{I}_1,\tilde{I}_k)\right) \\ &\le Cm_{\mathbf{n}}^N M \sum_{i=1}^{\infty} i^{N-1}\varphi(iq_{\mathbf{n}}) \\ &\le C\widehat{\mathbf{n}}q_{\mathbf{n}}^{-N\delta} \sum_{i=1}^{\infty} i^{N-1-N\delta}, \end{aligned}$$

by (6). This last tends to zero by the fact that $\widehat{\mathbf{n}}q_{\mathbf{n}}^{-N\delta} \to 0$ (see (H5)).

Proof of (12) : We have

$$Q_2 \equiv \widehat{\mathbf{n}}^{-1}E\Big(\sum_{i=2}^{2^N} T(\mathbf{n},x,i)\Big)^2 = \widehat{\mathbf{n}}^{-1}\left(\sum_{i=2}^{2^N} E\left[T(\mathbf{n},x,i)\right]^2 + \sum_{\substack{i,j=2,\ldots,2^N\\ i\neq j}} E\left[T(\mathbf{n},x,i)\right]\left[T(\mathbf{n},x,j)\right]\right).$$

By Cauchy-Schwartz inequality, we get :

$$\forall\ 2\le i\le 2^N : \widehat{\mathbf{n}}^{-1}E\left[T(\mathbf{n},x,i)\right]\left[T(\mathbf{n},x,j)\right] \le \left(\widehat{\mathbf{n}}^{-1}E\left[T(\mathbf{n},x,i)\right]^2\right)^{1/2}\left(\widehat{\mathbf{n}}^{-1}E\left[T(\mathbf{n},x,j)\right]^2\right)^{1/2}.$$

Then, it suffices to prove that

$$\widehat{\mathbf{n}}^{-1}E\left[T(\mathbf{n},x,i)\right]^2 \to 0 \quad ;\forall\ 2\le i\le 2^N.$$

We will prove this for $i=2$, the case where $i\neq 2$ is similar. We have $T(\mathbf{n},x,2) = \sum_{\mathbf{j}\in\mathcal{J}} U(2,\mathbf{n},x,\mathbf{j}) = \sum_{j=1}^{M}\hat{U}_j$, where we enumerate the $U(2,\mathbf{n},x,\mathbf{j})$ in the arbitrary way $\hat{U}_1,\ldots,\hat{U}_M$.

Then :

$$\begin{aligned}E\left[T(\mathbf{n},x,2)\right]^2 &= \sum_{i=1}^{M} Var\left(\hat{U}_i\right) + \sum_{i=1}^{M}\sum_{\substack{j=1\\ i\neq j}}^{M} Cov\left(\hat{U}_i,\hat{U}_j\right)\\ &= A_1 + A_2.\end{aligned}$$

The stationarity of the process $(X_{\mathbf{i}},Y_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^N}$, implies that :

$$\begin{aligned}Var\left(\hat{U}_i\right) &= Var\left(\sum_{\substack{i_k=1\\ k=1,\dots,N-1}}^{m_{\mathbf{n}}}\ \sum_{i_N=1}^{q_{\mathbf{n}}}\Delta_{\mathbf{i}}\right)^2\\ &= m_{\mathbf{n}}^{N-1}q_{\mathbf{n}}\ Var\left(\Delta_{\mathbf{i}}\right)\\ &\quad + \sum_{\substack{i_k=1\\ k=1,\dots,N-1}}^{m_{\mathbf{n}}}\ \sum_{i_N=1}^{q_{\mathbf{n}}}\ \underset{\mathbf{i}\neq\mathbf{j}}{\sum_{\substack{j_k=1\\ k=1,\dots,N-1}}^{m_{\mathbf{n}}}}\ \sum_{j_N=1}^{q_{\mathbf{n}}} E\Delta_{\mathbf{i}}\Delta_{\mathbf{j}}.\end{aligned}$$

We proved above that $Var\left(\Delta_{\mathbf{i}}\right) < C$. By Lemma 4.5.2, we have :

$$|E\Delta_{\mathbf{i}}(x)\Delta_{\mathbf{j}}(x)| \leq C\phi_x(a_{\mathbf{n}})^{-1}\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right). \tag{15}$$

Then, we deduce that

$$\begin{aligned}Var\left(\hat{U}_i\right) &\leq Cm_{\mathbf{n}}^{N-1}q_{\mathbf{n}}\left(1+\phi_x(a_{\mathbf{n}})^{-1}\sum_{\substack{i_k=1\\ k=1,\dots,N-1}}^{m_{\mathbf{n}}}\ \sum_{i_N=1}^{q_{\mathbf{n}}}\left(\varphi\left(\|\mathbf{i}\|\right)\right)\right)\\ &\leq Cm_{\mathbf{n}}^{N-1}q_{\mathbf{n}}\phi_x(a_{\mathbf{n}})^{-1}\sum_{\substack{i_k=1\\ k=1,\dots,N-1}}^{m_{\mathbf{n}}}\ \sum_{i_N=1}^{q_{\mathbf{n}}}\left(\varphi(\|\mathbf{i}\|\right).\end{aligned}$$

Consequently, we have :

$$A_1 \leq CMm_{\mathbf{n}}^{N-1}q_{\mathbf{n}}\phi_x(a_{\mathbf{n}})^{-1}\sum_{i=1}^{\infty} i^{N-1}\left(\varphi(i)\right).$$

Let

$$\begin{aligned}I\left(2,\mathbf{n},x,\mathbf{j}\right) &= \Big\{\mathbf{i} : j_k(m_{\mathbf{n}}+q_{\mathbf{n}})+1 \leq i_k \leq j_k(m_{\mathbf{n}}+q_{\mathbf{n}})+m_{\mathbf{n}},\quad 1\leq k\leq N-1;\\ &\qquad +j_N(m_{\mathbf{n}}+q_{\mathbf{n}})+m_{\mathbf{n}}+1 \leq i_N \leq (J_N+1)(m_{\mathbf{n}}+q_{\mathbf{n}})\Big\}.\end{aligned}$$

The variable $U(2,\mathbf{n},x,\mathbf{j})$ is the sum of the $\Delta_{\mathbf{i}}$ such that $\mathbf{i}$ is in $I(2,\mathbf{n},x,\mathbf{j})$. Since $m_{\mathbf{n}}>q_{\mathbf{n}}$, if $\mathbf{i}$ and $\mathbf{i}'$ are respectively in the two different sets $I(2,\mathbf{n},x,\mathbf{j})$ and $I(2,\mathbf{n},x,\mathbf{j}')$; then $i_k \neq i'_k$ for un certain k such that $1\le k\le N$ and $\|\mathbf{i}-\mathbf{i}'\|>q_{\mathbf{n}}$.
By using the definition of A_2, the stationarity of the process and (15), we have :

$$A_2 \le \sum_{\substack{j_k=1\\k=1,\ldots,N}}^{n_k} \sum_{\substack{i_k=1\\k=1,\ldots,N\\\|\mathbf{i}-\mathbf{j}\|>q_{\mathbf{n}}}}^{n_k} E\Delta_{\mathbf{i}}\Delta_{\mathbf{j}} \le C\phi_x(a_{\mathbf{n}})^{-1}\widehat{\mathbf{n}} \sum_{\substack{i_k=1\\k=1,\ldots,N\\\|\mathbf{i}\|>q_{\mathbf{n}}}}^{n_k} (\varphi(\|\mathbf{i}\|))$$

and

$$A_2 \le C\phi_x(a_{\mathbf{n}})^{-1}\widehat{\mathbf{n}}\sum_{i=q_{\mathbf{n}}}^{\infty} i^{N-1}\left(\varphi(i)\right).$$

We deduce that :

$$\begin{aligned}\widehat{\mathbf{n}}^{-1}E\left[T(\mathbf{n},x,2)\right]^2 &\le CMm_{\mathbf{n}}^{N-1}q_{\mathbf{n}}\widehat{\mathbf{n}}^{-1}\phi_x(a_{\mathbf{n}})^{-1}\sum_{i=1}^{\infty} i^{N-1-\delta}\\ &\quad +C\phi_x(a_{\mathbf{n}})^{-1}\sum_{i=q_{\mathbf{n}}}^{\infty} i^{N-1-\delta}.\end{aligned}$$

From $(m_{\mathbf{n}}+q_{\mathbf{n}})^{-N}m_{\mathbf{n}}^{N-1}q_{\mathbf{n}} = (m_{\mathbf{n}}+q_{\mathbf{n}})^{-N}m_{\mathbf{n}}^{N}\left(\frac{q_{\mathbf{n}}}{m_{\mathbf{n}}}\right)\le \frac{q_{\mathbf{n}}}{m_{\mathbf{n}}}$, we get :

$$\begin{aligned}CMm_{\mathbf{n}}^{N-1}q_{\mathbf{n}}\widehat{\mathbf{n}}^{-1}\phi_x(a_{\mathbf{n}})^{-1} &= \widehat{\mathbf{n}}(m_{\mathbf{n}}+q_{\mathbf{n}})^{-N}m_{\mathbf{n}}^{N-1}q_{\mathbf{n}}\widehat{\mathbf{n}}^{-1}\phi_x(a_{\mathbf{n}})^{-1}\\ &\le \left(\frac{q_{\mathbf{n}}}{m_{\mathbf{n}}}\right)\phi_x(a_{\mathbf{n}})^{-1}\\ &= q_{\mathbf{n}}s_{\mathbf{n}}\left(\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}})\right)^{\frac{-1}{2N}}\phi_x(a_{\mathbf{n}})^{-1}\\ &= q_{\mathbf{n}}s_{\mathbf{n}}\left(\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}})^{(1+2N)}\right)^{\frac{-1}{2N}}.\end{aligned}$$

By the hypothesis on $q_{\mathbf{n}}s_{\mathbf{n}}$, this last term converges to $\to 0$. Finally, we have :

$$C\phi_x(a_{\mathbf{n}})^{-1}\sum_{i=q_{\mathbf{n}}}^{\infty} i^{N-1-\delta} \le C\phi_x(a_{\mathbf{n}})^{-1}\int_{q_{\mathbf{n}}}^{\infty} t^{N-1-\delta}dt = C\phi_x(a_{\mathbf{n}})^{-1}q_{\mathbf{n}}^{N-\delta}.$$

This last term converges to zero by (8) and ends the proof of (12).
Proof of (13) : Let us use the following decomposition of small and big blocks

$$S'_{\mathbf{n}} = T(\mathbf{n},x,1) \qquad\qquad S''_{\mathbf{n}} = \sum_{i=2}^{2^N} T(\mathbf{n},x,i).$$

Then, we can write :

$$\widehat{\mathbf{n}}^{-1}E\left(S'_{\mathbf{n}}\right)^2=\widehat{\mathbf{n}}^{-1}ES^2_{\mathbf{n}}+\widehat{\mathbf{n}}^{-1}E\left(S''_{\mathbf{n}}\right)^2-2\widehat{\mathbf{n}}^{-1}ES_{\mathbf{n}}S''_{\mathbf{n}}.$$

Lemma 4.4.6(iii) and (12) imply respectively that $\widehat{\mathbf{n}}^{-1}E\left(S_{\mathbf{n}}\right)^2=\widehat{\mathbf{n}}^{-1}var\left(S_{\mathbf{n}}\right)\to\alpha(1-\alpha)$ and $\widehat{\mathbf{n}}^{-1}E\left(S''_{\mathbf{n}}\right)^2\to 0$.
Then, to show that $\widehat{\mathbf{n}}^{-1}E\left(S'_{\mathbf{n}}\right)^2\to\alpha(1-\alpha)$, it suffices to remark that $\widehat{\mathbf{n}}^{-1}ES_{\mathbf{n}}S''_{\mathbf{n}}\to 0$ because, by Cauchy-Schwartz's inequality, we can write :

$$\left|\widehat{\mathbf{n}}^{-1}ES_{\mathbf{n}}S''_{\mathbf{n}}\right|\le\widehat{\mathbf{n}}^{-1}E\left|S_{\mathbf{n}}S''_{\mathbf{n}}\right|\le\left(\widehat{\mathbf{n}}^{-1}ES^2_{\mathbf{n}}\right)^{1/2}\left(\widehat{\mathbf{n}}^{-1}ES''^2_{\mathbf{n}}\right)^{1/2}.$$

Recall that $T\left(\mathbf{n},x,1\right)=\sum_{\mathbf{j}\in\mathcal{J}}U\left(1,\mathbf{n},x,\mathbf{j}\right)$, so

$$\begin{aligned}\widehat{\mathbf{n}}^{-1}E\left(S'_{\mathbf{n}}\right)^2 &= \hat{\mathbf{n}}^{-1}\sum_{\substack{j_k=0\\k=1,\ldots,N}}^{r_k-1}E\left[U\left(1,\mathbf{n},x,\mathbf{j}\right)\right]^2\\ &+ \hat{\mathbf{n}}^{-1}\times\sum_{\substack{j_k=0\\k=1,\ldots,N}}^{r_k-1}\ \sum_{\substack{i_k=0\\k=1,\ldots,N\\ i_k\neq j_k \text{ for some k}}}^{r_k-1}cov[U(1,\mathbf{n},x,\mathbf{j}),U(1,\mathbf{n},x,\mathbf{i})]\end{aligned}$$

By similar arguments used above for A_2, this last term is not greater than

$$\begin{aligned}&C\phi_x(a_{\mathbf{n}})^{-1}\sum_{\substack{i_k=1\\k=1,\ldots,N\\\|\mathbf{i}\|>q_{\mathbf{n}}}}^{r_k-1}(\varphi(\|\mathbf{i}\|))\\ &\le C\phi_x(a_{\mathbf{n}})^{-1}\sum_{i=q_{\mathbf{n}}}^{\infty}i^{N-1}\left(\varphi(i)\right)\le C\phi_x(a_{\mathbf{n}})^{-1}q_{\mathbf{n}}^{N-\delta}\to 0.\end{aligned}$$

So $Q_3\to\alpha(1-\alpha)$. This ends the proof.

Proof of 14 : Since $|\Delta_{\mathbf{i}}|\le C\phi_x(a_{\mathbf{n}})^{-1/2}$, we have : $|U\left(1,\mathbf{n},x,\mathbf{j}\right)|\le Cm^N_{\mathbf{n}}\phi_x(a_{\mathbf{n}})^{-1/2}$. Then we deduce that

$$Q_4\le Cm^{2N}_{\mathbf{n}}\phi_x(a_{\mathbf{n}})^{-1}\widehat{\mathbf{n}}^{-1}\sum_{\substack{j_k=0\\k=1,\ldots,N}}^{r_k-1}P\left[\left|U\left(1,\mathbf{n},x,\mathbf{j}\right)\right|>\epsilon\left(\alpha(1-\alpha)\widehat{\mathbf{n}}\right)^{1/2}\right].$$

We have $\left|U\left(1,\mathbf{n},x,\mathbf{j}\right)\right|/\left(\left(\alpha(1-\alpha)\widehat{\mathbf{n}}\right)^{1/2}\right)\le Cm^N_{\mathbf{n}}\left(\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}})\right)^{-1/2}=C\left(s_{\mathbf{n}}\right)^{-N}\to 0$, because $m_{\mathbf{n}}=\left[\left(\widehat{\mathbf{n}}\phi_x(a_{\mathbf{n}})\right)^{1/(2N)}/s_{\mathbf{n}}\right]$ and $s_{\mathbf{n}}\to\infty$.

So, for all ϵ and $\mathbf{j} \in \mathcal{J}$; if $\widehat{\mathbf{n}}$ is great enough, then $P\left[U(1,\mathbf{n},x,\mathbf{j}) > \epsilon\left(\alpha(1-\alpha)\widehat{\mathbf{n}}\right)^{1/2}\right] = 0$. Then $Q_4 = 0$ for $\widehat{\mathbf{n}}$ great enough. This yields the proof.

Proof of Lemma 4.4.8

By change of variables, using the stationarity of the process, we have

$$\begin{aligned}
E\left[\alpha\hat{F}_D^x - \hat{F}_N^x(q_\alpha)\right] &= \alpha - \frac{1}{EK_{\mathbf{i}}}E\left[K_{\mathbf{i}}H_{\mathbf{i}}(q_\alpha)\right]\\
&= \alpha - \frac{1}{EK_{\mathbf{i}}}EK_{\mathbf{i}}E\left[H_{\mathbf{i}}(q_\alpha)\,|X_{\mathbf{i}}\right]\\
&= \alpha - \frac{1}{EK_{\mathbf{i}}}E\left(K_{\mathbf{i}}\int_{\mathbb{R}}K_2\left(\frac{q_\alpha-y}{b_{\mathbf{n}}}\right)f^{X_{\mathbf{i}}}(y)dy\right)\\
&= \alpha - \frac{1}{EK_{\mathbf{i}}}E\left(K_{\mathbf{i}}\int_{\mathbb{R}}b_{\mathbf{n}}{}^{(-1)}K_2^{(1)}\left(\frac{q_\alpha-y}{b_{\mathbf{n}}}\right)F^{X_{\mathbf{i}}}(y)dy\right)\\
&= \alpha - \frac{1}{EK_{\mathbf{i}}}E\left(K_{\mathbf{i}}\int_{\mathbb{R}}K_2^{(1)}(t)\,F^{X_{\mathbf{i}}}(q_\alpha - b_{\mathbf{n}}t)\,dt\right)\\
&= \alpha + \beta_1 + \beta_2
\end{aligned}$$

where

$$\beta_1 = -\frac{1}{EK_{\mathbf{i}}}E\left(K_{\mathbf{i}}\int_{\mathbb{R}}K_2^{(1)}(t)\,F^{X_{\mathbf{i}}}(q_\alpha)\,dt\right) = -\frac{1}{EK_{\mathbf{i}}}E\left(K_1\left(\frac{d(x,X)}{a_{\mathbf{n}}}\right)F^X(q_\alpha)\right)$$

and

$$\begin{aligned}
\beta_2 &= \frac{1}{EK_{\mathbf{i}}}E\left(K_{\mathbf{i}}\int_{\mathbb{R}}K_2^{(1)}(t)\left[F^{X_{\mathbf{i}}}(q_\alpha) - F^{X_{\mathbf{i}}}(q_\alpha - b_{\mathbf{n}}t)\right]dt\right)\\
&\leq \frac{1}{EK_{\mathbf{i}}}E\left(K_{\mathbf{i}}\int_{\mathbb{R}}K_2^{(1)}(t)\left|F^{X_{\mathbf{i}}}(q_\alpha) - F^{X_{\mathbf{i}}}(q_\alpha - b_{\mathbf{n}}t)\right|dt\right)\\
&\leq C\frac{1}{EK_{\mathbf{i}}}E\left(K_{\mathbf{i}}(b_{\mathbf{n}})^{b_2}\int_{\mathbb{R}}|t|^{b_2}K_2^{(1)}(t)\right)dt\\
&\leq C(b_{\mathbf{n}})^{b_2}.
\end{aligned}$$

This yields the proof of the first result of the Lemma. The following result ends the proof of the second result

$$\begin{aligned}
\beta_1 &= \frac{1}{EK_{\mathbf{i}}}E\left(K_1\left(\frac{d(x,X)}{a_{\mathbf{n}}}\right)\left[F^x(q_\alpha) - F^X(q_\alpha)\right]\right) - \frac{1}{EK_{\mathbf{i}}}E\left(K_1\left(\frac{d(x,X)}{a_{\mathbf{n}}}\right)F^x(q_\alpha)\right)\\
&\leq \frac{1}{EK_{\mathbf{i}}}E\left(K_1\left(\frac{d(x,X)}{a_{\mathbf{n}}}\right)\left|F^x(q_\alpha) - F^X(q_\alpha)\right|\right) - \alpha\\
&\leq C(a_{\mathbf{n}})^{b_1}\frac{1}{EK_{\mathbf{i}}}E\left(K_1\left(\frac{d(x,X)}{a_{\mathbf{n}}}\right)\right) - \alpha\\
&\leq C(a_{\mathbf{n}})^{b_1} - \alpha.
\end{aligned}$$

References

Abdi, A, Abdi, S, Dabo-Niang, S and Diop Aliou (2010a). $P-$mean consistency of a nonparametric conditional quantile estimator for random fields, *Mathematical Methods and Statistics*, 2010, Vol. 19, No. 1, 1-21..

Abdi, A, Abdi, S, Diop Aliou and Dabo-Niang, S(2010b). *Asymptotic normality of a nonparametric conditional quantile estimator for random fields*, Preprint.

Anselin, L. and Florax, R.J.G.M. (1995). *New Directions in Spatial Econometrics*, Springer, Berlin.

Biau, G. and Cadre, B. (2004). Nonparametric Spatial Prediction, *Stat. Inference Stochastic Proc.*, **Vol. 7**, pp. 327-349.

Carbon, M., Hallin, M. and Tran, L.T. (1996). Kernel density estimation for random fields, *Stat. and Probab. Lett.*, **Vol. 36**, pp. 115-125.

Carbon, M., Tran, L.T. and Wu, B. (1997). Kernel density estimation for random fields : the L_1 theory, *Nonparam. Stat.*, **Vol. 6**, pp. 157-170.

Cressie, N.A.C. (1991). *Statistics for spatial Data*, Wiley Series in Probability and Mathematical Statistics, New-York.

Dabo-Niang, S. and Yao, A.F. (2007). Density estimation for spatial functional random variables, *Preprint.*

Dabo-Niang, S. and Rhomari, N.(2003) Estimation non paramétrique de la régression avec variable explicative dans un espace métrique. *Comptes. Rendus. Acad. Sci. Paris*, Ser I. 336, 75-80.

Dabo-Niang, S. and Laksaci, A., (2007), *Nonparametric estimation of conditional quantiles when the regressor is valued in a semi-metric space*, submitted.

Dabo-Niang, S. and Laksaci, A., (2009), *Nonparametric quantile regression estimation for functional dependent data*, submitted.

Dabo-Niang, S et Thiam, B. (2010). L_1 consistency of a kernel estimate of spatial conditional quantile. *Statistics and probability letters*, Vol. 80 (17-18), 1447-1458.

Ferraty, F. and Vieu, Ph. (2006), *Nonparametric functional data analysis.* Springer-Verlag, New-York.

Guyon, X., 1987. Estimation d'un champ par pseudo-vraisemblance conditionnelle : Etude asymptotique et application au cas Markovien. Proceedings of the Sixth Franco-Belgian Meeting of Statisticia

Guyon, X. (1995). *Random Fields on a Network - Modeling, Statistics, and Applications*, Springer, New-York.

Hallin, M., Lu, Z. and Tran, L.T. (2004a). Kernel density estimation for spatial processes : the L_1 theory, *J. of Mult. Anal.*, **Vol. 88**, **No. 1**, pp. 61-75.

Hallin, M., Lu, Z. and Tran, L.T. (2004b). Local linea r spatial regression, *Ann. Stat.*, **Vol. 32**, **No. 6**, pp. 2469-2500.

Hallin, M., Lu, Z. and Yu, K. (2009). Local linear spatial quantile regression, *Bernouilli,* **vol. 15**, Number 3, pp. 659-686.

Honda T. (2000), Nonparametric estimation of a conditional quantile for a- mixing processes, *Annals of the Institute of Statistical Mathematics*, **52**, pp. 459- 470.

Koenker, R. (2000). Galton, Edgrworth, Frish and prospect for quantile regression in econometrics. *Journal of Econometrics.*, **95**, pp. 347-374.

Koenker, R. (2005). *Quantile Regression* Cambridge University Press in econometrics. Cambridge, U.K.

Koenker, R. and Mizera, I. (2004). Penalized triograms : total variation regularization for bivariate smoothing. *Journal of the Royal Statistical Society Series B*, **66**, pp. 145-164.

Koul, H. L. and Mukherjee, K. (1994). Regression quantiles and related processes under longrange dependence, *Journal of Multivariate Analysis*, **51**, pp. 318- 337.

Lu, Z. and Chen, X. (2002). Spatial nonparametric regression estimation : Non-isotropic Case, *Acta Mathematicae Applicatae Sinica, English Series*, **Vol. 18, No.4**, pp. 641-656.

Lu, Z. and Chen, X. (2004). Spatial kernel regression estimation : weak concistency, *Stat. and Probab. Lett.*, **Vol. 68**, pp. 125-136.

Lu, Z. (1996). Weak consistency of nonparametric kernel regression under alpha-mixing dependence. *Chinese Science Bulletin*, **Vol. 41**, pp. 2219-2221.

Lu, Z. and Chen, P. (1997). Distribution-free strong consistency for nonparametric kernel regression involving nonlinear time series, *J. of Stat. Planning. and Inference,* **Vol. 65**, pp. 67-86.

Portnoy, S. L. (1991). Asymptotic behavior of regression quantiles in nonstationary dependent cases. *Journal of Multivariate Analysis*, **38**, pp. 100-113.

Rio, E. (2000). *Théorie Asymptotic des Processus Aléatoires Faiblement Dépendants,* Springer-Verlag, Berlin.

Ripley, B. (1981). *Spatial Statistics*, Wiley, New-York.

Roussas, G.G. (1988). Nonparametric estimation in mixing sequences of random variables. *J. of Stat. Plann. and Infer.*, **Vol. 18**, pp. 135-149.

Stone, C. J., (1977), *Consistent nonparametric regression. Discussion.* Ann. Stat. **5**, 595-645.

Tran, L.T. (1990). Kernel density estimation on random fields, *J. of Mult. Anal.*, **Vol. 34**, pp. 37-53.

Tran, L.T. and Yakowitz, S. (1993). Nearest neighbor estimators for random fields, *J. of Mult. Ana*l., **Vol. 44**, pp. 23-46.

Yu, K., Lu, Zudi. and Stander, J., (2003), *Quantile regression : applications and current research areas*, The Statistician. **52**, 3, 331-350.

Chapitre 5

Applications et Discussions

Le cadre général de cette contribution est la modélisation non paramétrique des données fonctionnelles spatialement dépendantes. D'après la définition de Delicado et al. (2008), la statistique fonctionnelle spatiale étudie les phénomènes dont les observations $(Z_{\mathbf{s}})_{\mathbf{s}\in S}$ sont des variables aléatoires indexées par un sous ensemble S dans $\mathbb{R}^d$, Z_s appartenant à un espace d'état fonctionnel. Cette thématique a été sélectionnée par Ramsay (2008) parmi les huit plus importants sujets de la statistique moderne. De même que pour la statistique spatiale vectorielle, on peut distinguer d'un point de vie pratique trois types d'observations fonctionnelles spatialement dépendantes : celles aisément obtenues à partir d'un champ aléatoire à temps continu, ou d'un processus spatio-temporel, ou de processus ponctuels spatiaux. Dans le reste de ce paragraphe, on liste quelques exemples d'applications de données spatiales pour lesquels nos modèles conditionnels jouent un rôle important

5.1 Sur la mise en oeuvre des modèles étudiés en géostatistique

L'analyse des données géostatistiques concerne le cas où le champ aléatoire fonctionnel est obtenu à partir d'un champ aléatoire $(Z_{\mathbf{s}})_{\mathbf{s}\in S}$ où S est un sous-espace continu de R^d. En pratique, on trouvera ce type de données dans différents domaines concernant par exemple les ressources miniers, pétroliers ou agronomiques. A ce stade, on peut appliquer nos modèles (voir Section 3.4 du Chapitre 3) pour étudier les caractéristiques d'un terrain géographique. En effet, la base de données **soil250** du package **geoR** de **R** est un exemple concret traitant les propriétés chimiques d'un échantillon de sol. Spécifiquement, on peut prendre comme exemple $Z_{s\in[0,45]\times[0,120]}$ comme étant le PH d'un sol observé sur une région. Sur les graphes suivants on représente les réalisations de ce champ aléatoire Z_s

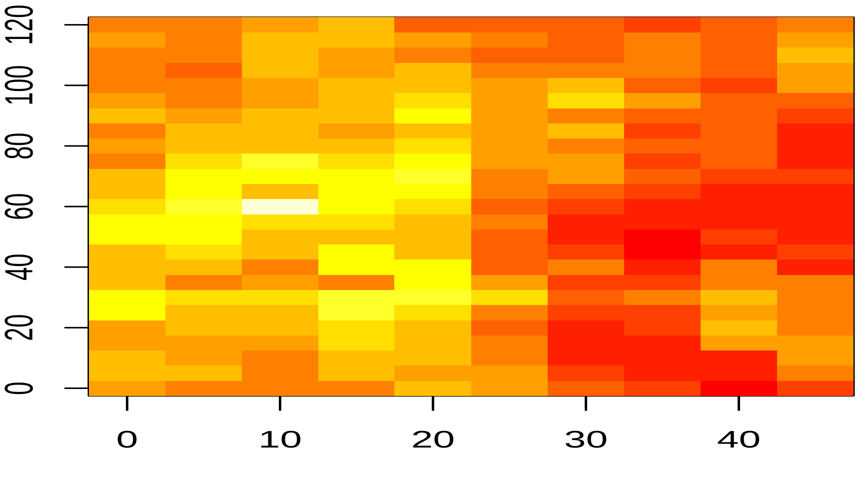
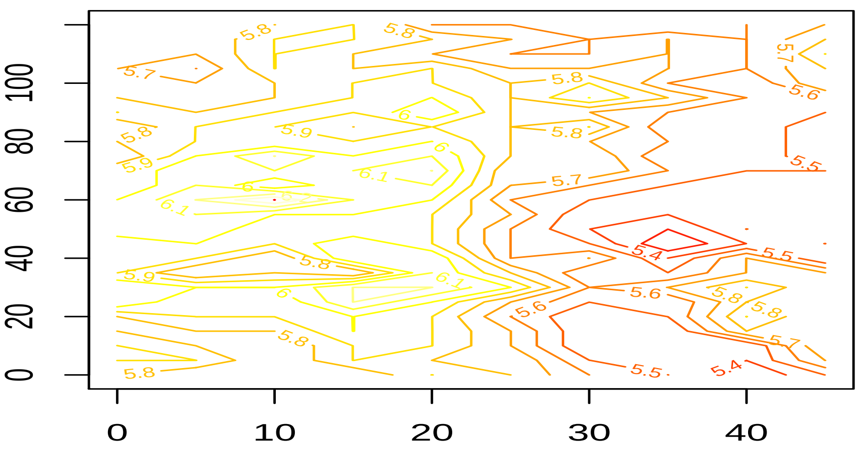


Image du PH d'un sol dans une région donnée.



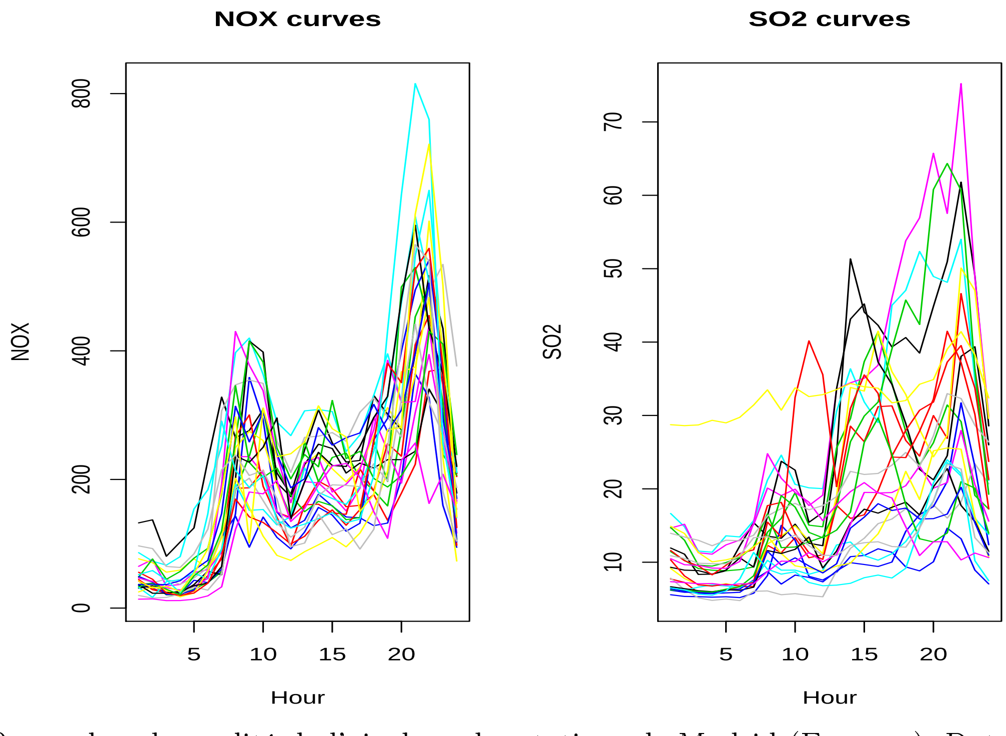
Contour-plot du PH

On garde les mêmes notations que celles de la Section 3.4 du Chapitre 3 et on sélectionne aléatoirement une collection de sites $\mathcal{G}_{\mathbf{n}}$ de cardinal finie $\widehat{\mathbf{n}}$. Ainsi, pour prévoir la quantité du PH sur un site s_0, on suppose que cette caractéristique dépend seulement des quantités de PH dans le voisinage noté $\mathcal{V}_{\mathbf{s_0}}$. En utilisant ce voisinage on peut construire des variables

fonctionnelles spatialement dépendantes sur les sites d'observations t_i dans $\mathcal{G}_{\mathbf{n}}$ par les quantités suivantes $X_{t_i} = (Z_t,\, t \in \mathcal{V}_{t_i})$ où $\mathcal{V}_{\mathbf{t}_k} = \mathcal{V} + \mathbf{t}_k$ avec $\mathcal{V} = \mathcal{V}_{\mathbf{s_0}} - \mathbf{s_0}$. Par conséquent, on peut prendre comme approximation de Z_{s_0} (la quantité du PH sur un site $\mathbf{s_0}$) le mode estimé $\hat{\theta}(X_{\mathbf{s_0}})$ (ou le quantile $\hat{q}_{.5}(X_{\mathbf{s_0}})$ qui est obtenu à l'aide des observations $(X_{\mathbf{t_k}}, Y_{\mathbf{t_k}})_{k=1,\ldots,\widehat{\mathbf{n}}}$ où $Y_{\mathbf{t_k}} = Z_{\mathbf{t_k}}$.
Notons que, la prévision de la quantité du PH en des sites de la région où les données ne sont pas disponibles donne des indications importantes sur la qualité du sol de ces sites, en particulier les caractéristiques hydro-dynamiques (la circulation de l'eau et de l'air) qui est un facteur principal pour le rendement d'une récolte.

5.2 Applicabilité sur des données latticielles

Les données latticielles sont des observations d'un champ aléatoire $(Z_{\mathbf{s}})_{\mathbf{s}\in S}$ où S est une collection dénombrable de points de R^d. C'est le cas par exemple lorsque on s'intéresse à étudier des données observées sur des stations espacées ou des zones administratives ayant un réseau. A titre d'exemple, on considère les données de pollution suivantes :



NOX et SO_2 courbes de qualité de l'air dans des stations de Madrid (Espagne). Date=18/01/2009.

Ces données représentent les courbes journalières des NOx, SO_2 sur 24 stations situées à Madrid (Espagne). Les données correspondent à la journée du 18 Janvier 2009. En utilisant le mode conditionnel et/ou la médiane conditionnelle on peut prévoir les pics de concentration d'ozone dans une station sachant les courbes journalières des NOx, SO_2 dans les autres stations voisines. En effet, on prend comme variable fonctionnelle $X_{\mathbf{i}}$ la courbe de NOx (resp. de $S0_2$) sur la station **i** livrée sur une grille de discrétisation de 24 points représentant les heures de la journée. La variable réponse est $Y_{\mathbf{i}} = \max_{t=0,1,\ldots 23}(O_3(t))$ sur la station **i**. Donc, les quantités

$$\hat{\theta}(X_{\mathbf{i_0}}) = \arg\max \widehat{f}^{X_{\mathbf{i_0}}}(y) \text{ ou } \hat{q}_{.5}(X_{\mathbf{i_0}}) = \inf\{y| \quad \widehat{F}^{X_{\mathbf{i_0}}}(y) \geq .5\}$$

où

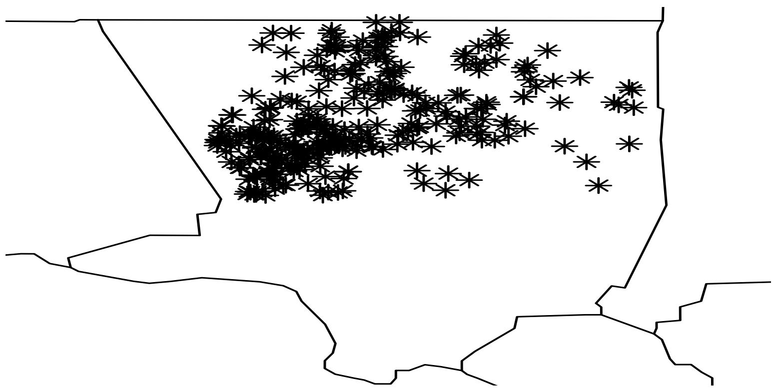
$$\widehat{f}^x(y) = \frac{h_H^{-1}\sum_{\mathbf{i}\neq\mathbf{i_0}} K(h_K^{-1}d(x, X_{\mathbf{i}}))H(h_H^{-1}(y - Y_{\mathbf{i}}))}{\sum_{\mathbf{i}\neq\mathbf{i_0}} K(h_K^{-1}d(x, X_{\mathbf{i}}))},$$

$$\widehat{F}^x(y) = \frac{\sum_{\mathbf{i}\neq\mathbf{i_0}} K(h_K^{-1}d(x, X_{\mathbf{i}}))H(h_H^{-1}(y - Y_{\mathbf{i}}))}{\sum_{\mathbf{i}\neq\mathbf{i_0}K}(h_K^{-1}d(x, X_{\mathbf{i}}))}, \qquad \forall y \in \mathbb{R}$$

peuvent être de bonnes approximations du pics de concentration d'ozone dans la station $\mathbf{i_0}$ sachant la courbe de NOx (resp. de SO_2) dans la même station $\mathbf{i_0}$. Notons que ce problème de prévision est très utile en pratique. Il nous permet de déterminer les gaz les plus polluants. Ce qui est important pour connaître les causes de la pollution afin de proposer des règlements pour contrôler la qualité de l'aire.

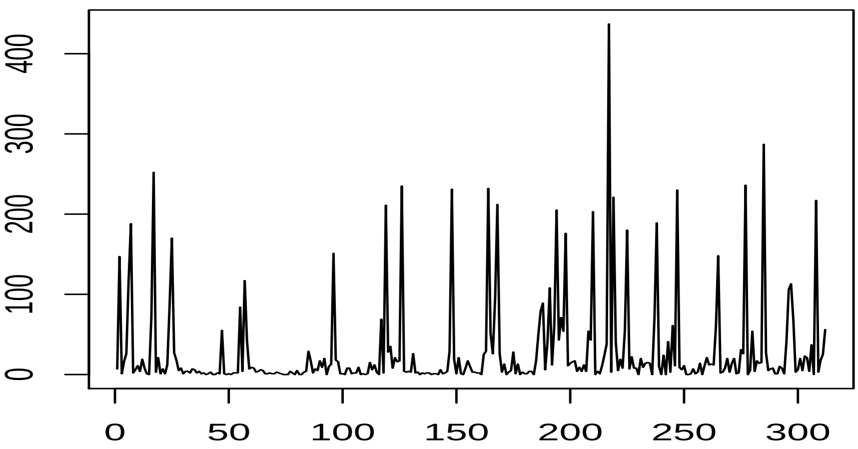
5.3 Applicabilité sur des données spatio-temporelles

L'étude des données spatio-temporelles est un cas particulier de l'analyse des données géostatistiques. Plus précisément, elle concerne le cas où le domaine S est un sous ensemble de $\mathbb{R}^2 \times R^+$ (espace et temps). Néanmoins, l'analyse de ce type de données demande un traitement assez particulier car pour la plupart, des cas la dépendance sur l'espace est différente de celle du temps. Autrement dit, on doit faire face à deux dépendances séparées, une sur l'espace et une autre sur le temps. On trouvera dans la littérature un nombre important d'exemples de données spatio-temporelles, on citera par exemple le cas de données d'incendie de forêts en Californie dans le Nord de Los Angeles (la base des données **fires** du package **spacetime** de **R**) pour la période 1976- 2000. Le graphe suivant donne les positions géographiques de ces incendies.



Les positions géographiques des incendies

Tandis que la courbe courbe suivante représente le temps entre deux incendies



Le temps entre deux incendies

Il est clair qu'avec ce jeu de donnés, nous disposons des positions géographiques s_i et des temps t_i des incendies. A partir de ces données on peut construire un champ aléatoire fonctionnel en utilisant les réalisations du processus spatio-temporel décrivant le temps entre

deux incendies consécutifs c'est-à-dire des observations $Z_{(s_i,t_j)} = t^{s_i}_{j+1} - t^{s_i}_j$, où $t^{s_i}_j$ est le jième temps d'observation de l'incendie d'un site s_i.
En reprenant les mêmes arguments que précédemment, on peut choisir aléatoirement un ensemble de points $\mathcal{G}_{\mathbf{n}}$ et supposer que la valeur de (Z_{s_0,t_0}) dépend seulement d'observations dans le voisinage $\mathcal{V}_{\mathbf{s_0}} \times \mathcal{V}_{\mathbf{t_0}}$. En utilisant $\mathcal{V}_{\mathbf{s_0}} \times \mathcal{V}_{\mathbf{t_0}}$, on peut construire les variables fonctionnelles spatialement dépendantes de la même manière que le cas des données géostatistiques comme suit $X_{s_i,t_i} = (Z_{s,t},\, (s,t) \in \mathcal{V}_{s_i,t_i})$ où $\mathcal{V}_{s_k,t_k} = \mathcal{V}_{s_k} \times \mathcal{V}_{t_k} = \mathcal{V} + (s_k, t_K)$ avec $\mathcal{V} = \mathcal{V}_{s_0} \times \mathcal{V}_{t_0} - (s_0, t_0)$. Les quantités $\hat{\theta}(X_{s_0,t_0})$ et $\hat{q}_{.5}(X_{s_0,t_0})$ peuvent être de bonnes approximations de (Z_{s_0,t_0}) en utilisant la valeur de Z continûment observée sur une zone géographique $\mathcal{V}_{s_0}$ et un intervalle de temps limité $\mathcal{V}_{t_0}$.

Chapitre 6

Commentaires et perspectives

L'objectif de ce chapitre est de montrer l'étendue du sujet étudié en discutant son implication dans d'autres problèmes statistiques très importants aussi bien en pratique qu'en théorie. Nous insistons sur le fait que l'étude de ces problèmes se situe directement en continuité des travaux effectués dans le cadre de cette thèse qui prennent l'avantage sur la régression classique dans certains cas.

6.1 Sur d'autres outils de prévision spatiale

Dans plusieurs situations, la prévision via les régions prédictives est beaucoup plus adéquate que la prévision ponctuelle. A ce sujet, les résultats obtenus dans cette thèse pourraient être employés pour adapter des régions et/ou des intervalles prédictives largement connus en série chronologique. En effet, gardons les notations des chapitres précédents et fixons un $\alpha \in (0,1)$.

- *MCD-région (Maximum conditional density region)*
  La MCD région prédictive a été introduite par Hyndman (1995), elle est définie par

$$MCD_\alpha(x) = \{y, \quad f^x(y) \geq C_\alpha(x)\}$$

où

$$C_\alpha(x) = \max\{C > 0;\ P\left[\{y,\, f^x(y) \geq C\}|X = x\right] \geq \alpha\}.$$

Il est clair que la convergence presque complète du deuxième chapitre sur le mode conditionnel est une conséquence directe de la convergence presque complète de la densité conditionnelle. Ainsi, on peut utiliser ce résultat comme un support théorique justifiant l'approximation suivante

$$\widehat{MCD}_\alpha(x) = \left\{y, \quad \widehat{f}^x(y) \geq \widehat{C_\alpha}(x)\right\}$$

où

$$\widehat{C_\alpha}(x) = \max\left\{C > 0;\ \widehat{P}\left[\{y,\, \widehat{f}^x(y) \geq C\}|X = x\right] \geq \alpha\right\}$$

$\widehat{P}$ est l'estimateur spatial de la probabilité conditionnelle déterminée par $\widehat{F}^x$ ou $\widehat{f}^x$. Nous renvoyons à De Gooijer et Gannoun (2000) pour d'autres définitions équivalentes. Il est indiqué dans ce dernier article que cet ensemble est le plus petit au sens de la mesure de Lebesgue parmi toutes les régions prédictives de seuil α.

- *SCM-Intervalle (shortest conditional modal interval)*
Cette notion fut proposée par Lientz (1970) pour le cas d'une distribution non conditionnelle. Le cas conditionnel a été abordé par De Gooijer et Gannoun (2000). Dans cet article, ce concept est donné via la fonction de répartition conditionnelle par l'intervalle
$$[a_{SCM_\alpha}(x), b_{SCM_\alpha}(x)] = \arg\min \{Leb[a,b], \quad F^x(b) - F^x(a) \geq \alpha\}$$
où $Leb[a,b]$ désigne la mesure de Lebesgue de l'intervalle $[a,b]$.
De même que le cas précédent, l'intervalle SCM est le plus petit intervalle parmi tous les intervalles prédictifs de seuil α. La détermination pratique de cet intervalle est basée sur l'estimation de la fonction de répartition conditionnelle. Cette dernière a été traitée comme modèle préliminaire à l'estimation des quantiles conditionnels du chapitre 4. Plus précisément, on considère l'approximation
$$[a_{SCM_\alpha}(x), b_{SCM_\alpha}(x)] = \arg\min \left\{Leb[a,b], \quad \widehat{F}^x(b) - \widehat{F}^x(a) \geq \alpha\right\}$$
où
$$\widehat{F}^x_{\mathbf{n}}(y) = \frac{\sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} K_1\left(\frac{d(x,X_\mathbf{i})}{a_\mathbf{n}}\right) K_2\left(\frac{y-Y_\mathbf{i}}{b_\mathbf{n}}\right)}{\sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} K_1\left(\frac{d(x,X_\mathbf{i})}{a_\mathbf{n}}\right)}$$
Notons que pour la question du choix du paramètre de lissage, on peut utiliser le critère (8) du Chapitre 3 sur la densité conditionnelle. Bien que l'utilisation de ce critère est justifiée par le lien étroit entre les deux modèles (mode et quantile), son optimalité asymptotique reste une question ouverte.
- *CP-Intervalle (Conditional percentile interval)* Le CP-intervalle de seuil α est donné par
$$[q_\alpha(x), q_{1-\alpha}(x)].$$
Ainsi, la propriété asymptotique étudiée dans le chapitre 4 nous permet de prendre comme approximation l'intervalle suivant
$$[\hat{q}_\alpha(x), \hat{q}_{1-\alpha}(x)].$$
En ce qui concerne la détermination pratique de cet intervalle, il est nécessaire de développer une méthode de sélection automatique des paramètres intervenant dans le calcul des quantités $\hat{q}_\alpha(x)$ et $\hat{q}_{1-\alpha}(x)$, tels que les paramètres de lissage a_n, b_n et la métrique d. Pour cela, on propose d'utiliser la règle de sélection de Ferraty et Veu (2006) avec une pénalisation comme suit
$$CVCDF(a_n, b_n, d) = \sum_{\mathbf{l},\mathbf{l}\in I_\mathbf{n}} \left[1\!\!1_{Y_\mathbf{k}\leq Y_\mathbf{l}} - \widehat{F}^{X_\mathbf{k}}_{-\mathbf{k}}(Y_\mathbf{l}, a_n, b_n, d)\right]^2$$

$$\widehat{F}_{-\mathbf{k}}^{X_{\mathbf{k}}}(y, a_n, b_n, d) = \frac{\sum_{\mathbf{l} \in I^{\mathbf{k}}_{\mathbf{n},\varsigma_\mathbf{n}}} K(h_K^{-1} d(X_\mathbf{k}, X_\mathbf{l})) H(h_H^{-1}(y - Y_\mathbf{l}))}{\sum_{\mathbf{l} \in I^{\mathbf{k}}_{\mathbf{n},\varsigma_\mathbf{n}}} K(h_K^{-1} d(X_\mathbf{k}, X_\mathbf{l}))}$$

avec

$$I^{\mathbf{k}}_{\mathbf{n},\varsigma_\mathbf{n}} = \{\mathbf{l}; \quad \|\mathbf{l} - \mathbf{k}\| \geq \varsigma_\mathbf{n} \}.$$

Cependant, cette question du choix des paramètres de lissage et le choix de la métrique dans l'estimation des quantiles conditionnels reste toujours des problèmes ouverts en statistique fonctionnelle.

Finalement, nous rappelons que cette diversité des choix dans les régions prédictives est une particularité importante pour les modèles non paramétriques liés à la distribution conditionnelle. Contrairement à la régression classique où la seule région prédictive est obtenue par la normalité asymptotique.

6.2 Sur d'autres modèles conditionnels

Le sujet des modèles non paramétriques liés à la distribution conditionnelle est un thème très vaste et notre objectif ici n'est pas de présenter une discussion exhaustive de ces modèles. Nous allons concentrer notre discussion sur un modèle peu traité en statistique fonctionnelle, qui s'appelle E.S.(Expected shortfall) conditionnel. Ce modèle a été introduit par Acerbi (2002) comme mesure de risque en finance alternative à la VaR. Il a été récemment considéré en statistique fonctionnelle par Ferraty et al. (2010) en utilisant la distribution conditionnelle. Plus précisément, ils ont adopté la définition suivante

Definition 6.2.1 *(voir Ferraty et al.(2010) Soit un seuil $\alpha \in (0,1)$, le déficit prévu d'une variable réelle Y conditionnellement à une variable fonctionnelle X est la fonction*

$$CES_\alpha(x) = \alpha^{-1} \int_{VaR_\alpha}^{+\infty} t f^x(t) dt$$

où VaR_α et le quantile conditionnel d'ordre α très petit et f^x est la densité conditionnelle de Y sachant $X = x$.

Ce modèle non paramétrique est largement utilisé en finance en prenant Y une variable aléatoire représentant la rentabilité d'un portefeuille, la valeur d'une action ou d' un indice boursier. Le caractère spatio-fonctionnel des données financiers ou des marchés boursiers est un facteur primordial avec la mondialisation de l'économie et le développement des réseaux des communications et Internet. Il est naturel dans certains cas que l'estimation non paramétrique repose essentiellement sur celle de la densité conditionnelle f^x et du quantile conditionnel $q_\alpha(x)$. Typiquement, on estime le CES de seuil α par

$$\widehat{\mathrm{CES}}_\alpha = \alpha^{-1} \int_{\widehat{\mathrm{VaR}}_{\alpha(x)}}^{+\infty} t \widehat{f}^x(t) dt$$

où $\widehat{f}^x$ (resp.$\widehat{\mathrm{VAR}}_\alpha(x)$) est l'estimateur de la densité conditionnelle (resp. de quantile conditionnel d'ordre α). Les propriétés asymptotiques de cet estimateur du CES sont étroitement liées à ceux de l'estimateur de la densité conditionnelle ainsi que l'estimateur $\widehat{q}_\alpha(x)$. L'étude de ce modèle est un projet promoteur en perspective.

6.3 Sur la supériorité des modèles étudiés par rapport à la régression

Bien que la régression classique est le thème le plus privilégiée en statistique non paramétrique, ce dernier est très limité pour certaines situations. A titre d'exemple, en statistique classique lorsque la densité conditionnelle $f^x(\cdot)$ est de type

$$f^x(y) = \frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{1}{2}(y-x)^2\right\} + \exp\left\{-\frac{1}{2}(y+x)^2\right\}\right),$$

il est connu que la régression classique ne marche pas bien. Cet exemple est souvent présenté en statistique vectorielle pour argumenter la supériorité du mode conditionnel par rapport à la régression. Dans ce contexte de la statistique fonctionnelle plusieurs études comparatives entre la régression et les autres modèles conditionnels ont été développés (voir par exemple Ferraty et al (2006), Attouch et al (2009)). Le travail présenté dans cette thèse montre que ces modèles non paramétriques conditionnels ne sont pas des simples outils de prévision alternatifs à la régression, mais sont aussi des modèles d'analyse statistique exploitable dans d'autres problèmes statistiques. En effet, dans la première section de ce chapitre nous avons donné une variante des régions prédictives pour lesquelles nos modèles conditionnels jouent un rôle important pour les constructions de ces régions. Par contre, les seuls intervalles prédictifs qu'on peut obtenir par la régression classique sont ceux de la normalité asymptotique de son estimateur. Notons que, la gestion du risque est un autre domaine de la statistique pour laquelle la régression classique est moins importante que les autres modèles conditionnels. En effet, la régression classique modélise les évènements centraux, or la question principale dans la modélisation de risque est de prévenir les évènements rares qui causent un risque majeur. Il est clair que les quantiles sont des modèles adéquats pour répondre à cette problèmatique de risque.

En conclusion, on peut dire que le domaine d'application des modèles conditionnels considérés dans cette thèse est beaucoup plus vaste que celle de la régression classique. Par ailleurs, ce dernier est moins robuste que les autres modèles conditionnels dans plusieurs situations. Toutefois, la régression reste toujours le modèle le plus simple à manipuler en prévision. Ainsi, il serait souhaitable de chercher comment combiner de façon efficace tous ces modèles afin d'élaborer une méthode de prévision compatible dans toute situation. Le développement d'un tel mélange automatique facile à utiliser est un sujet de recherche très prometteur.

6.4 Sur le test de la dépendance spatiale en pratique

On revient dans cette section à la question du Chapitre 2 sur la quantification de la dépendance spatiale en pratique. Il est évident que ce problème a des influences directes sur l'efficacité de l'analyse statistique considérée. Une mauvaise exploration de la corrélation des données dans l'échantillonnage entraîne une perte flagrante de la qualité de l'estimation. Ce problème a été largement abordé en statistique spatial classique. On dispose de deux principaux outils de test de la dépendance spatiale tel l'indice de Moran ou l'indice de Geary (voir Guyon, 2008). Pour le cas spatial fonctionnel, la question reste toujours ouverte. Cependant, comme indiqué dans le deuxième chapitre, la plupart des auteurs intègre la dépendance spatiale dans le calcul de l'estimateur en introduisant l'ensemble de contiguïté comme suit

$$\widehat{f}^{X_{\mathbf{k}}}(y, W^{\mathbf{n}}_{\mathbf{k}}) = \frac{h_H^{-1} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K(h_K^{-1} d(X_{\mathbf{k}}, X_{\mathbf{i}})) H(h_H^{-1}(y - Y_{\mathbf{i}})) 1\!\!1_{W^{\mathbf{n}}_{\mathbf{k}}}(\mathbf{i})}{\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K(h_K^{-1} d(X_{\mathbf{k}}, X_{\mathbf{i}})) 1\!\!1_{W^{\mathbf{n}}_{\mathbf{k}}}(\mathbf{i})}$$

où $W^{\mathbf{n}}_{\cdot}$ est un sous ensemble défini par une matrice de poids $W = (w_{\mathbf{i},\mathbf{k}})$ par

$$W^{\mathbf{n}}_{\mathbf{k}} = \{\mathbf{i}, \text{ tel que } \ w_{\mathbf{i},\mathbf{k}} \leq \iota_{\mathbf{n}}\}.$$

Bien que cette procédure montre une grande compatibilité en pratique (voir Dabo-Niang et al (2011)) mais son optimalité au sens que

$$\exists W^{0,\mathbf{n}}_{\mathbf{k}} \quad \text{tel que} \quad \frac{\mathbb{E}[\widehat{f}^{X_{\mathbf{k}}}(y, W^{0,\mathbf{n}}_{\mathbf{k}}) - f^{X_{\mathbf{k}}}(y)]^2}{\inf_{W^{\mathbf{n}}_{\mathbf{k}}} \mathbb{E}[\widehat{f}^{X_{\mathbf{k}}}(y, W^{\mathbf{n}}_{\mathbf{k}}) - f^{X_{\mathbf{k}}}(y)]^2} \to 1$$

reste toujours un problème ouvert. Il est clair que l'optimalité de cette procédure est étroitement liée à la matrice de poids W. En faisant une allusion avec le cas vectoriel on peut distinguer trois types de cas (voir Haining, 2004)

– *Le cas où les données sont issues d'un champ aléatoire à indices continus*
  Dans ce cas, on peut utiliser la fonction trace-variogramme définie par $\gamma(\mathbf{k}, \mathbf{i}) = \frac{1}{2}\int Var[X_{\mathbf{k}} - X_{\mathbf{i}}](t)dt$ et estimée par

$$\widehat{\gamma}(\mathbf{l}, \mathbf{k}) = \frac{1}{2\# N_{\mathbf{l},\mathbf{k}}} \sum_{\mathbf{i},\mathbf{j}\in N_{\mathbf{l},\mathbf{k}}} d(X_{\mathbf{i}}, X_{\mathbf{j}})$$

  où $N_{\mathbf{l},\mathbf{k}} = \{\mathbf{i}, \mathbf{j} \in \mathcal{I}_{\mathbf{n}} \text{ tel que } \ \|\mathbf{i}-\mathbf{j}\| = \|\mathbf{l}-\mathbf{k}\|\}$ et d est la métrique de L^2. Nous renvoyons à (Giraldo, 2009) pour le code calculant cette fonction.
– *Le cas où le champ aléatoire est référencé par des pixels*
  Dans le cas où les données sont référencées en pixels, on prend comme matrice de poids $W = (w_{ij})$ où w_{ij} est le nombre de pas nécessaire pour aller d'un pixel $i = (i_p, i_q)$ à un pixel $j = (j_p, j_q)$.

– *Le cas des données latticielles*
  Dans ce cas où les observations sont $Z_{\mathbf{i}}$ ou $\mathbf{i}$ désigne une zone pré-définie, une station d'observations ou une unité administrative, on peut définir notre matrice $w_{\mathbf{ij}}$ comme une fonction de la distance $d_{\mathbf{ij}}$ entre les sites $\mathbf{i}$ et $\mathbf{j}$ ou bien une fonction de la longueur des frontières, comme on peut la prendre comme fonction de n'importe quelle interaction entre $\mathbf{i}$ et $\mathbf{j}$ telle que l'intensité des réseaux de communication entre $\mathbf{i}$ et $\mathbf{j}$ si on est en zone urbaine, ...

Notons que les matrices de poids définies ci-dessus peuvent être utilisées pour construire une version fonctionnelle de l'indice de Moran ou de Geary.
Alternativement, à cette procédure basée sur les ensembles de contiguïté, on peut incorporer la corrélation spatiale dans la pondération. Autrement dit, on estime nos modèles par

$$\widehat{f}^{X_{\mathbf{i}}}(y) = \sum_{\mathbf{j}\neq\mathbf{i}} W(X_{\mathbf{i}}, X_{\mathbf{j}}, \mathbf{i}, \mathbf{j}) H(h_H^{-1}(y - Y_{\mathbf{j}})) \quad \text{pour la densité conditionnelle}$$

$$\widehat{F}^{X_{\mathbf{i}}}(y) = \sum_{\mathbf{j}\neq\mathbf{i}} W(X_{\mathbf{i}}, X_{\mathbf{j}}, \mathbf{i}, \mathbf{j}) K_2(h_H^{-1}(y - Y_{\mathbf{j}})) \quad \text{pour la fonction de répartition conditionnelle}$$

où $W(X_{\mathbf{i}}, X_{\mathbf{j}}, \mathbf{i}, \mathbf{j})$ sont des fonctions positives vérifiant

$$\sum_{\mathbf{j}\neq\mathbf{i}} W(X_{\mathbf{i}}, X_{\mathbf{j}}, \mathbf{i}, \mathbf{j}) = 1.$$

De même que le cas précédent la quantification de la dépendance spatiale est définie selon le cas étudié. A titre d'exemple Muller (2008) considère pour la régression locale linéaire, une pondération qui est une fonction de la distance entre les sites. Il est à noter que certains auteurs (voir Hamdad et al., 2011) considèrent un lissage sur les sites en prenant une pondération de type

$$W(X_{\mathbf{i}}, X_{\mathbf{j}}) = \frac{K(h_K^{-1} d(X_{\mathbf{i}}, X_{\mathbf{j}}) L(h_L^{-1}(\|\mathbf{i}-\mathbf{j}\|))}{\sum_{\mathbf{j}\neq\mathbf{i}} K(h_K^{-1} d(X_{\mathbf{i}}, X_{\mathbf{j}}) L(h_L^{-1}(\|\mathbf{i}-\mathbf{j}\|))}$$

L est un noyau et h_L est une suite de paramètres réels.
En conclusion, on peut dire qu'il y a plusieurs façons d'exploiter le caractère spatial des données et que le choix approprié de la méthode est lié à la nature des données (géostatistique ou latticielle,), type de mixing (polynomiale ou exponentielle), la stationnarité (isotropique ou non),.... Cependant, le choix optimal est un très bon projet en perspective.

6.5 Sur la convergence uniforme et d'autres méthodes d'estimation

– **La convergence uniforme en statistique fonctionnelle spatiale.** La convergence uniforme en statistique fonctionnelle est un sujet très important, il a été initié par

Ferraty et al. (2010) . Nous renvoyons à cette référence pour plus de discussions sur l'impact et l'originalité de ce problème en statistique fonctionnelle. L'obtention de ce résultat pour nos modèles conditionnels spatiaux ouvre une grande parenthèse vers l'adaption des méthodes *data-driven* en statistique fonctionnelle spatiale.

- **L'estimation robuste des quantiles conditionnels** Le traitement des quantiles conditionnels comme modèle non paramétrique robuste est une approche alternative pour estimer les quantiles conditionnels. Cette méthode d'estimation a été abondamment étudiée en statistique vectorielle. Nous renvoyons à Laksaci et al (2008, et 2011) pour le cas fonctionnel et Dabo-Niang et Thiam (2010) pour le cas spatial. Nous espérons étendre les résultats de ces trois derniers travaux au cas spatio-fonctionnel.
- **L'estimation par la méthode locale linéaire pour le mode conditionnel** Une généralisation naturelle de notre méthode d'estimation (méthode du noyau) appelée la méthode locale constante est la méthode locale linéaire. Récemment, Demongeot et al.(2012) ont étudié l'estimation de la densité conditionnelle par cette méthode. Ils ont établi dans le cas i.i.d, la convergence presque complète de l'estimateur construit et ils ont appliqué leur résultat à l'estimation du mode conditionnel. Les outils développés dans cette thèse devraient parvenir à s'adapter à cette méthode d'estimation.
- **L'estimation par la méthode locale linéaire pour les quantiles conditionnels** Dans cette thèse nous traitons la fonction des quantiles conditionnels comme inverse de la fonction de répartition conditionnelle. Cependant, cette considération n'a pas lieu lorsqu'on utilise la méthode locale linéaire, car on ne peut garantir l'inversiblité de l'estimateur. Ainsi, les idées de Demongeot et al.(2012) ne sont pas automatiquement applicables pour le cas des quantiles conditionnels. Il devrait être possible de combiner les idées de Yu et Jones (1998) avec ceux de Xu et Wang (2008) pour étudier les propriétés asymptotiques de l'estimateur local linéaire de quantiles conditionnels spatiaux.

Chapitre 7

Bibliographie générale

Abdi, A, Abdi, S, Dabo-Niang, S and Diop Aliou (2010a). $P-$mean consistency of a nonparametric conditional quantile estimator for random fields, *Mathematical Methods and Statistics*, 2010, **19**, 1-21.

Abdi, A, Abdi, S, Diop Aliou and Dabo-Niang, S(2010b). *Asymptotic normality of a nonparametric conditional quantile estimator for random fields*, Preprint.

Acerbi, C (2002). Spectral measures of risk : A coherent representation of subjective risk aversion, *Journal of Banking and Finance*, **26** (7), 1505-1518.

Anselin, L. and Florax, R.J.G.M. (1995). *New Directions in Spatial Econometrics*, Springer, Berlin.

Attouch, M. Laksaci, A. and Ould Saïd E (2009). Asymptotic distribution of robust estimator for functional nonparametric models. *Comm. Statist. Theory and Methods* **38**, 1317-1335.

Attouch, M. Laksaci, A. and Ould Saïd E. (2010). Asymptotic normality of a robust estimator of the regression function for functional time series data. *J. of the Korean Statist. Soc.*, **39**, 489-500.

Attouch, M., Chouaf, B. and Laksaci, A. (2012) Nonparametric M-estimation for functional spatial data , *Communication of the Korean Statistical society,* **19**, 193-211

Attouch, M.; Gheriballah A. et A. Laksaci Convergence Presque complète d'un estimateur robuste de la régression non paramétrique fonctionnelle : Cas spatial, *Pub. Inst. Stat. Univ. Paris LI*, **56**. 3-16 .

Azzedine, N. Laksaci, A. and Ould Saïd E. (2008). On the robust nonparametric regression estimation for functional regressor. *Statist. Probab. Lett.* **78**, 3216-3221

Barrientos-Marin, J., Ferraty, F. and sc Vieu, P. (2010). Locally Modelled Regression and Functional Data. *Journal of Nonparametric Statistics*, **22**, 617-632

Bàillo, A. and Grané, A. (2009). Local linear regression for functional predictor and scalar response, *Journal of Multivariate Analysis*, **100**, Pages 102–111

Barrientos-Marin, J., Ferraty, F. and sc Vieu, P. (2010). Locally Modelled Regression and

Functional Data. *Journal of Nonparametric Statistics*, **22**, 617-632

Basse, M. ; Dabo-Niang, S. ; Diop, A. (2008) Mean square properties of a class of kernel density estimates for spatial functional random variables. *Ann. I.S.U.P.* **52**, 91–108.

Benhenni, K., Griche-Hedli, S. and Rachdi, M. (2010). Estimation of the regression operator from functional fixed-design with correlated errors. *Journal of Multivariate Analysis*, **101**, 476–490.

Berlinet, A., Gannoun, A. and Matzner-Lober, E. (1998a). Normalité asymptotique d'estimateurs convergents du mode conditionnel. *La Rev. Canad. de Statist.*, **26**, 365-380.

Berlinet, A., Gannoun, A., Matzner-Løber, E., (1998b), *Propriétés asymptotiques d'estimateurs convergents des quantiles conditionnels.* C. R. Acad. Sci., Paris, Sér. I, Math. **326**, 611-614.

Berlinet, A. ; Elamine, A. ; Mas, A. (2011), Local linear regression for functional data. *Ann. Inst. Statist. Math.* **63**, 1047–1075

Biau, G. (2003). Spatial Kernel Density Estimation, *Math. Meth. Stat.*, **12**, 371-390.

Biau, G. and Cadre, B. (2004). Nonparametric spatial prediction. *Stat. Inference Stoch. Process.* **7**, 327-349.

Bosq, D. (2000). Linear processes in function spaces. Theory and Applications, Lecture Notes in Statistics, **149**, Springer-Verlag.

Cadre, B. (2001). Convergent estimators for the L^1-median of a Banach valued random variable. *Statistics.* **35**, 509-521.

Carbon, M., Hallin, M. and Tran, L.T. (1996). Kernel density estimation for random fields, *Stat. and Probab. Lett.*, **36**, 115-125.

Carbon, M., Tran, L.T. and Wu, B. (1997). Kernel density estimation for random fields : the L_1 theory, *Nonparam. Stat.*, **6**, 157-170.

Carbon, M., Francq, C. and Tran, L. T.(2007). Kernel regression estimation for random fields. *J. Statist. Plann. Inference.* **137**, 778-798.

Crambes, C. Delsol, L. and Laksaci, A. (2008). Robust nonparametric estimation for functional data. *J. Nonparametric Statist.* **20**, 573-598.

Chouaf, A. and Laksaci, A. (2011). On the functional local linear estimate for spatial regression. *Statistics & Risk Modeling* **29**, 189-214.

Collomb, G., Härdle, W. and Hassani, S. (1987). A note on prediction via conditional mode estimation. *J. Statist. Plann. and Inf.*, **15**, 227-236.

Cressie, N. A. (1991). *Statistics for spatial data.* Wiley Series in Probability and Mathematical Statistics. New York.

Dabo-Niang, S. (2004). Kernel density estimator in an infinite dimensional space with a rate of convergence in the case of diffusion process. *Applied Math. Lett.*, **17**, 381-386.

Dabo-Niang, S. (2002). Estimation de la densité dans un espace de dimension infinie : Application aux diffusions. *C. R. Math. Acad. Sci. Paris* . **334**, 213-216.

Dabo-Niang, S. et Rhomari, N. (2003). Estimation non paramétrique de la régression avec variable explicative dans un espace métrique. *C. R., Math., Acad. Sci. Paris.* **336**, 75-80.

Dabo-Niang, S. and Yao, A-F. (2007). Kernel regression estimation for continuous spatial processes. *Math. Methods. Statist.* **16**, 298-317.

Dabo-Niang, S. et Laksaci, A. (2007). Propriétés asymptotiques d'un estimateur à noyau du mode conditionnel pour variable explicative fonctionnelle. *Ann. I.S.U.P.* **51** , 27-42.

Dabo-Niang, S., Rhomari, N. (2009) Kernel regression estimation in a Banach space. *J. Statist. Plann. Inference* **139**, 1421–1434.

Dabo-Niang, S, Yao, A-F, Pischedda, L, Cuny, P et Gilbert, F. (2010). Spatial mode estimation for functional random fields with application to bioturbation problem. *Stochastic Environmental Research and Risk Assessment,* **24,** 487-497

Dabo-Niang, S. and Laksaci, A. (2010). Note on conditional mode estimation for functional dependent data, *Statistica*, **70**, 83-94.

Dabo-Niang, S. and Thiam, B. (2010). Robust quantile estimation and prediction for spatial processes. *Statist. Probab. Lett.* **80**, 1447-1458.

Dabo-Niang, S., Rachdi, M. and Yao, A-F. (2011). Spatial kernel regression estimation and prediction for functional random fields. *Far. East. J. Statist.*, **37**, 77-113.

Dabo-Niang, S. and Laksaci, A. (2011a). Nonparametric quantile regression estimation for functional dependent data. Comm in Statist. Theory & Methods **41**, 1254-1268.

Dabo-Niang,S. Kaid,Z. Laksaci, A. (2011b) Sur la régression quantile pour variable explicative fonctionnelle : Cas des données spatiales *C. R., Math., Acad. Sci. Paris.* **349,** 1287-1291.

Dabo-Niang,S. Kaid,Z. Laksaci, A. (2012) On spatial conditional mode estimation for a functional regressor *Statist. Probab. Lett.* **82,** 1413-1421.

Dabo-Niang,S. Kaid,Z. Laksaci, A. (2012) Spatial conditional quantile regression : Weak consistency of a kernel estimate, *submitted.*

Dabo-Niang,S. Kaid,Z. Laksaci, A. (2012) Asymptotic properties of the kernel estimate of the spatial conditional mode when the regressor is functional, *preprint.* Delsol, L. (2007) CLT and $\mathbb{L}^q$ errors in nonparametric functional regression. *C. R. Math. Acad. Sci. Paris* **345** 411–414.

Delsol, L (2009). Advances on asymptotic normality in non-parametric functional time series analysis. *Statistics*, **43**, 13–33.

Demongeot, J., Laksaci, A., Madani, F. and Rachdi, M. (2011). Functional data : Local linear estimation of the conditional density and its application. *Statistics*, In presse.

Diggle, P., Ribeiro, P.J. (2007) *Model-Based Geostatistics.* Springer, New York.

El Machkouri, M. (2011) Asymptotic normality of the Parzen-Rosenblatt density estimator for strongly mixing random fields. it Stat. Inference Stoch. Process. **14** 73–84.

Ezzahrioui, M. and Ould-Saïd, E. (2008a). Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data. *J. Nonparametr. Stat.* **20**, 3-18.

Ezzahrioui, M. and Ould-Saïd, E. (2008b). Asymptotic normality of the kernel estimator of conditional quantiles in a normed space. *Far East J. Theor. Stat.* **25**, 15-38.

Ezzahrioui, M. and Ould-Saïd, E. (2008c). Asymptotic results of a nonparametric conditional quantile estimator for functional time series. *Comm. Statist. Theory Methods.* **37**, 2735-2759.

Ferraty, F., Vieu, P. (2000). Dimension fractale et estimation de la régression dans des espaces vectoriels semi-normés. C. R. Acad. Sci., Paris, **330**, 139-142.

Ferraty, F., Vieu, P. (2004). Nonparametric models for functional data, with application in regression times series prediction and curves discrimination. J. Nonparametric Statist., **16**, 111-127.

Ferraty, F., Laksaci, A. and Vieu, P. (2005). Functional times series prediction via conditional mode. *C. R. Acad. Sci. Maths. Paris*, **340**, 389–392.

Ferraty, F., Rabhi, A. and Vieu, P. (2005). Conditional quantiles for dependent functional data with application to the climatic El Niño phenomenon. *Sankhya.* **67** , 378-398.

Ferraty, F., Vieu, P. (2006). Nonparametric functional data analysis. Theory and Practice. Springer-Verlag.

Ferraty, F., Laksaci, A. and Vieu, P.(2006). Estimation some characteristics of the conditional distribution in nonparametric functional models. *Stat Inference Stoch. Process.* **9**, 47-76.

Ferraty, F. Mas, A. and Vieu, P. (2007). Advances in nonparametric regression for functional variables. *Aust. and New Zeal. J. of Statist.* **49**, 1-20.

Ferraty, F., Rabhi, A. et Vieu, P. (2008). Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle. *Rev. Roumaine Math. Pures Appl.* **53**, 1-18.

Ferraty, F., Laksaci, A., Tadj, A. and Vieu, P. (2010). Rate of uniform consistency for nonparametric estimates with functional variables. *J. Statist. Plann. Inference.* **140**,335-352.

Ferraty, F., Laksaci, A., Tadj, A. and Vieu, P. (2011) Kernel regression with functional response. *Electron. J. Stat.* **5** , 159–171.

Ferraty, F. and Romain, Y. (2011). *The Oxford handbook of functional data analysis.* Oxford University Press.

Ferraty, F., Van Keilegom, I., Vieu, P. (2012) Ression when both response and predictor are functions *J. Multivariate Anal*, **109,**, 10-28.

Ferraty, F., Laksaci, A., Tadj, A. and Vieu, P. (2012). Estimation de la fonction de régression pour variable explicative et réponse fonctionnelles dépendantes *C. R. Acad. Sci. Maths. Paris* (sous presse)

Gannoun, A., Saracco, J. and Yu,K. (2003) Nonparametric prediction by conditional median and quantiles,*J. Statist. Plann. Inference* **117**, 207–223.

Gao, J. ; Lu, Z. ; Tjøstheim, D. (2008) Moment inequalities for spatial processes. *Statist. Probab. Lett.* **78** 687–697.

Gasser, T. Hall, P. and Presnell, B. (1998). Nonparametric estimation of the mode of a distribution of random curves. *J. R. Stat. Soc.Ser. B, Stat. Methodol.* **60**, 681-691.

Gheriballah, A., Laksaci, A. and Rouane, R. (2010). Robust nonparametric estimation for spatial regression. *J. Statist. Plann. Inference.* **140**, 1656-1670.

Giraldo, R.(2009) Geostatistical Analysis of Functional Data. Ph.D. thesis. Universitat Politècnica de Catalunya.

Guyon, X. (1995). *Random Fields on a Network-Modeling, Statistics and applications.* Springer, New York.

Hallin, M., Lu, Z. and Tran, L.T. (2004). Local linear spatial regression. *Ann. Stat.* **32**, 2469-2500.

Hallin, M., Lu, Z. and Yu, K. (2009). Local linear spatial quantile regression, *Bernouilli.* **15**, 659-686.

Haining, R. (2003) Spatial Data Analysis : Theory and Practice. Cambridge Univ Press.

Honda T. (2000), Nonparametric estimation of a conditional quantile for a- mixing processes, *Annals of the Institute of Statistical Mathematics*, **52**, pp. 459- 470.

Ioannides, D. A. ; Matzner-Løber, E.(2002) Nonparametric estimation of the conditional mode with errors-in-variables : strong consistency for mixing processes. *J. Nonparametr. Stat.* **14** , 341–352

Ioannides, D. A. ; Matzner-Løber, E. (2004) A note on asymptotic normality of convergent estimates of the conditional mode with errors-in-variables. *J. Nonparametr. Stat.* **16** 515–524

Khardani, S. ; Lemdani, M. and Ould Saïd, E. (2010) Some asymptotic properties for a smooth kernel estimator of the conditional mode under random censorship. *J. Korean Statist. Soc.* **39** , 455–469.

Khardani, S. ; Lemdani, M. and Ould Saïd, E. (2011) Uniform rate of strong consistency for a smooth kernel estimator of the conditional mode for censored time series. *J. Statist. Plann. Inference* **141** , 3426–3436

Khardani, S. ; Lemdani, M. and Ould Saïd, E. (2012) On the strong uniform consistency of the mode estimator for censored time series. *Metrika* **75**, 229–241.

Koenker, R. (2000). Galton, Edgrworth, Frish and prospect for quantile regression in econometrics. *Journal of Econometrics.*, **95**, pp. 347-374.

Koenker, R. (2005). *Quantile Regression* Cambridge University Press in econometrics. Cambridge, U.K.

Koenker, R. and Mizera, I. (2004). Penalized triograms : total variation regularization for bivariate smoothing. *Journal of the Royal Statistical Society Series B*, **66**, pp. 145-164.

Koul, H. L. and Mukherjee, K. (1994). Regression quantiles and related processes under long range dependence, *Journal of Multivariate Analysis*, **51**, pp. 318- 337.

Laïb, N. and Louani, D. (2010). Nonparametric kernel reghression estimate for functional stationary ergodic data : Asymptotic properties. *J. Multivariate Anal.*, **101**, 2266-2281.

Laïb, N. and Louani, D. (2011). Strong consistency of the regression function estimator for functional stationary ergodic data *J. Statist. Plan. and Inference,* **141**, 359-372.

Laksaci, A., Yousfate, A. (2002), Estimation fonctionnelle de la densité de l'opérateur de transition d'un processus de Markov à temps discret *C. R., Math., Acad. Sci. Paris.* **334**, 1035-1038.

Laksaci, A. (2007). Erreur quadratique de l'estimateur à noyau de la densité conditionnelle à variable explicative fonctionnelle. *C. R. Math. Acad. Sci. Paris.* **345**, 171-175.

Laksaci, A. and Maref, F. (2009). Estimation non paramétrique de quantiles conditionnels pour des variables fonctionnelles spatialement dépendantes. *C. R., Math., Acad. Sci. Paris.* **347**, 1075-1080.

Laksaci, A., Lemdani, M., Ould Saïd, E. (2009) A generalized L^1-approach for a kernel estimator of conditional quantile with functional regressors : Consistency and asymptotic normality. *Statist. Probab. Lett.*, 79, 1065-1073.

Laksaci, A. ; Mechab, B. (2010) Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle : cas des données spatiales. *Rev. Roumaine Math. Pures Appl.* **55**, 35–51.

A. Laksaci, M. Lemdani and E. Ould-Saïd, (2011). Asymptotic results for an L^1-norm kernel estimator of the conditional quantile for functional dependent data with application to climatology. *Sankhya* **A 73** 125–141 .

Laksaci, A., Madani, F. and Rachdi, M. (2011). Kernel conditional density estimation when the regressor is valued in a semi-metric space *Commun. Stat., Theory and Methods,* to appear.

Lemdani, M. ; Ould-Saïd, E. ; Poulin, N. (2009) Asymptotic properties of a conditional quantile estimator with randomly truncated data. *J. Multivariate Anal.* **100**, 546–559.

Li, J. and Tran, L. T. (2007). Hazard rate estimation on random fields. *J. Multivariate Anal.* **98**, 1337-1355.

Li, J. and Tran, L. T. (2009). Nonparametric estimation of conditional expectation. *J. Statist. Plann. Inference.* **139**, 164-175. Liang, H. de Uña-Álvarez, J. (2011) Asymptotic properties of conditional quantile estimator for censored dependent observations. *Ann. Inst. Statist. Math.* **63** , 267–289

Liang, H. de Uña-Álvarez, J. (2010) Asymptotic normality for estimator of conditional mode under left-truncated and dependent observations. *Metrika* **72**, 1–19.

Lin,Z. and Li, D.(2007) Asymptotic normality for L_1-norm kernel estimator of conditional median under association dependence,*J. Multivariate Anal.* **98** , 1214–1230.

Lu, Z. (1996). Weak consistency of nonparametric kernel regression under alpha-mixing dependence. *Chinese Science Bulletin*, **41**, pp. 2219-2221.

Lu, Z. and Chen, P. (1997). Distribution-free strong consistency for nonparametric kernel regression involving nonlinear time series, *J. of Stat. Planning. and Inference,* **65**, pp. 67-86.

Lu, Z. and Chen, X. (2002). Spatial nonparametric regression estimation : Non-isotropic Case, *Acta Mathematicae Applicatae Sinica, English Series*, **18**, pp. 641-656.

Lu, Z. and Chen, X. (2004). Spatial kernel regression estimation : weak concistency, *Stat. and Probab. Lett.*, **68**, pp. 125-136.

Masry, E. (2005). Nonparametric regression estimation for dependent functional data : Asymptotic normality. Stoch. Proc. and their Appl., **115**, 155-177.

Louani, D. and Ould-Saïd, E., (1999), Asymptotic normality of kernel estimators of the conditional mode under strong mixing hypothesis. *J. Nonparametric Stat.*, **11**, 413-442.

Nakhapetyan, B.S. (1987) An approach to the proof of limit theorems for dependent random variables. *Theory of Probab. and its Appl.*, 32, 535-539.

Nerini, D. ; Monestiez, P. ; Manté, C. (2010) Cokriging for spatial functional data. it J. Multivariate Anal. **101**, 409–418.

Ould-Saïd, E., (1997), A note on ergodic processes prediction via estimation of the conditional mode function. *Scand. J. Stat.* **24**, 231-239.

Ould Abdi, A. ; Diop, A. ; Dabo-Niang, S. ; Ould Abdi, S.A. (2010b) Estimation non paramétrique du mode conditionnel dans le cas spatial. *C. R. Math. Acad. Sci. Paris* **348**, 815–819.

Ould Saïd, E., Djabrane Y., (2011) Asymptotic normality of a kernel conditional quantile estimator under strong mixing hypothesis and left-truncation. *Comm. Statist. Theory Methods* **40**, 2605–2627

Ould Saïd, E.and Tatachak, A. (2011) A nonparametric conditional mode estimate under RLT model and strong mixing condition. it Int. J. Stat. Econ. **6** , 76–92.

Quintela del Rio, A. and Vieu, Ph., (1997), A nonparametric conditionnal mode estimate. *Nonparametric Statistics*, **8**, 253-266.

Portnoy, S. L. (1991). Asymptotic behavior of regression quantiles in nonstationary dependent cases. *Journal of Multivariate Analysis*, **38**, 100-113.

Quintela-del-Río, A. (2008). Hazard function given a functional variable : Non-parametric estimation under strong mixing conditions.*J. Nonparametr. Stat.* **20**, 413-430.

Rachdi, M., Vieu, P. (2007). Nonparametric regression functional data : Automatic smoothing parameter selection. J. Statist. Plan. Inf., **137**, 2784-2801.

Ramsay, J. O., Silverman, B. W. (1997). Functional data analysis. Springer, New York.

Ramsay, J. O. and Silverman, B. W. (2002) *Applied functional data analysis ; Methods and case studies.* Springer-Verlag, New York.

Ramsay, J.O., Silverman, B.W. (2005) *Functional Data Analysis, Second Edition.* Springer, New York.

Ramsay, J. (2008). Fda problems that I like to talk about. Personal communication.

Ripley, B. (1981). *Spatial Statistics.* Wiley, New York.

Roussas, G. (1969) Nonparametric estimation of the transition distribution function of a Markov process. *Ann. Math. Statist.* **40** 1386–1400.

Rio, E. (2000). *Théorie Asymptotic des Processus Aléatoires Faiblement Dépendants,* Springer-Verlag, Berlin.

Roussas, G. (1988). Nonparametric estimation in mixing sequences of random variables. *J. of Stat. Plann. and Infer.*, **18**, 135-149.

Roussas, G. (1991) Estimation of transition distribution function and its quantiles in Markov processes : strong consistency and asymptotic normality. it Nonparametric functional estimation and related topics , (Spetses, 1990), 443–462.

Robinson, P. M. (2011) Asymptotic theory for nonparametric regression with spatial data. *J. Econometrics* **165**, 5–19

Samanta, M., (1989), Non-parametric estimation of conditional quantiles. *Stat. Probab . Lett..* **7**, 407-412.

Samanta, M. ; Thavaneswaran, A. (1990) Nonparametric estimation of the conditional mode. *Comm. Statist. Theory Methods*, **19** , no. 12, 4515–4524

Sarda, P. and Vieu, P. (2000). Kernel regression. In M. Schimek (ed.) *Smoothing and regression ; Approaches, Computation, and Application.*, Wiley Series in Probability and Statistics, Wiley, New-York.

Stone, C. J., (1977), Consistent nonparametric regression. Discussion. *Ann. Stat.* **5**, 595-645.

Stute, W. (1986) On almost sure convergence of conditional empirical distribution functions. *Ann. Probab.* **14** 891–901.

Tran, L. T. (1990). Kernel density estimation on random fields.*J. Mult. Anal.* **34**, 37-53.

Tran, L.T. and Yakowitz, S. (1993). Nearest neighbor estimators for random fields, *J. of Mult. Anal*., **44**, 23-46.

Xu, R. and Wang, J. (2008). L_1-estimation for spatial nonparametric regression. *J. Nonparametr. Stat.* **20**, 523-573.

Youndjé, E., (1993), Estimation non paramétrique de la densité conditionnelle par la méthode du noyau. Thèse de Doctarat, Université de Rouen.

Yu, K., Lu, Zudi. and Stander, J., (2003), Quantile regression : applications and current research areas, *The Statistician.* **52**, 3, 331-350

Zhou,Y. and Liang, H. (2000) Asymptotic normality for L_1-norm kernel estimator of conditional median under α-mixing dependence, *J. Multivariate Anal.* **73** 136–154.